CONTENIDO

N.° 64

Estudios de Filosofía

Objetivos

La revista *Estudios de Filosofía* busca servir de instrumento de comunicación de los resultados parciales o finales de trabajos de investigación. Ser instrumento de divulgación de estudios de filosofía, tanto sobre problemas concernientes a la historia de la filosofía como a las discusiones contemporáneas en torno a problemas relativos a todos los campos de la investigación filosófica. Fomentar la difusión de los estudios de filosofía en la forma de producción original, comentarios, réplicas, reseñas o traducciones, promoviendo el intercambio de ideas y opiniones, y la discusión entre los autores colaboradores.

Información editorial

Los autores interesados en la publicación de sus trabajos en *Estudios de Filosofía* pueden consultar la sección nota a los autores en esta misma revista, visitar la página web https://revistas.udea.edu.co/index.php/estudios_de_filosofia, o escribir a la dirección electrónica: revistafilosofia@udea.edu.co

Los artículos de esta revista aparecen registrados en los siguientes índices:

- **Publindex** - Índice Nacional de Publicaciones Científicas**,** Colciencias, Colombia
- **SciELO** - Citation Index
- **Latindex.** Sistema Regional de Información en Línea para Revistas Científicas de América Latina, El Caribe, España y Portugal
- **Philosopher's Index.** Philosophy Documentation Center, Estados Unidos
- **Ulrich's Periodicals Directory,** Estados Unidos
- **Dialnet,** España
- **EBSCO** - Fuente Académica Premier
- **REDALYC** - Red de Revistas Científicas de América Latina y el Caribe, España y Portugal
- **ERIH PLUS** - European Reference Index for the Humanities and Social Sciences
- **IBZ** - Internationale Bibliographie der Geistes
- **DOAJ** - Directory of Open Access Journals
- **REDIB** - Red Iberoamericana de Innovación y Conocimiento Científico
- **Philapapers**
- **J-Gate**
- **V-lex**

Estudios de Filosofía

ISSN 0121-3628 | ISSN-e: 2256-358X | Julio-diciembre de 2021

Memory and perception, insights at the interface: editors' introduction*

Abstract: The recent development of specialized research fields in philosophy of memory and philosophy of perception invites a dialogue about the relationship between these mental capacities. Following a brief review of some of the key issues that can be raised at the interface of memory and perception, this introduction provides an overview of the contributions to the special issue, and outlines possible directions for further research.

1. The relation between memory and perception: inviting a dialogue

Memory and perception have been fundamental topics since the beginning of philosophy. Indeed, one of the central principles of Western philosophy has been the notion of the primacy of perception. Perception is often thought to be the most basic cognitive act, the act from which many other cognitive capacities are thought to derive, or to derive their content, none more so than episodic memory (Casey, 1991/2004, p. 137). But what, precisely is the relation between memory and perception?

The recent development of specialized research fields in philosophy of memory and philosophy of perception invites a dialogue about the relationship between these mental capacities. There are signs that people are starting to take an interest in this area: in February 2021, the Centre for Philosophy of Memory organized a workshop, *Memory and perception: starting the conversation,* which focused precisely on the relation between memory and perception. Yet, there is also evidence that there is lots

---

of work to be done: to highlight an important example, *The Routledge Handbook of Philosophy of Memory* (Bernecker & Michaelian, 2017), one of the key texts in the field, doesn't include a chapter on memory and perception. The motivation for this special issue is precisely to encourage more work and continue the conversation about the relation between memory and perception.

The intersection between the philosophy of memory and the philosophy of perception is a potentially fruitful domain of research. There are many questions that can be raised about the interface of memory and perception, many connections or divergences discovered. The question of how the content of memory relates to the content of perception is a fundamental concern. Is episodic memory the preservation of past perceptual content? The causal theory of memory ensures a tight link between perceiving and remembering, positing that genuine memories are appropriately causally connected to events that were perceived in the past through a continuously stored memory trace (Martin & Deutscher, 1966; Robins, 2016).

Yet many of the representations of episodic memory typically diverge somewhat from one's original experience. For example, one interesting phenomenon to consider is observer perspectives in episodic memory. Such memories are viewed from-the-outside, and one sees oneself in the remembered scene (Debus, 2007; Sutton, 2010). An interesting question is whether such memories cannot be said to authentically represent past perceptual experience. Many theorists think that there is something distorted about such memories (De Brigard, 2014; Fernández, 2015). Yet others think that, despite the detached point of view, observer perspectives can authentically represent past perceptual experience (McCarroll, 2018; *Cf*. Bernecker, 2015).[1] Thinking about observer perspectives, and other ways in which memory appears to be constructed rather than reproduced, is one of the interesting issues that arise when one considers the relationship between memory and perception. In fact, the evidence that episodic memory is an inherently creative and constructive process, which is tightly connected to imaginative processes (Addis, 2018), would seem to call into question the preservationist view of remembering. On a simulationist view of memory, remembering is a form of imagining (Michaelian, 2016; *Cf*. De Brigard, 2014; Shanton & Goldman, 2010). If remembering is a process more akin to imagining, indeed perhaps even just a form of imagining, does it still require a causal connection to the past perceptual experience? If remembering is an imaginative process, what is the relation between the content of perception and memory?

Another key question relates to the phenomenology of memory and perception. Do memory and perception share some phenomenal properties, or are they inherently distinct? Perception is often thought to involve a feeling of presence (Nanay, 2018). What you perceive is phenomenally present, it feels present both temporally and spatially. Yet

---

1   See also section 3 below on the distinction between truth and authenticity in episodic memory.

is the same true of remembering? Episodic memory is sometimes characterized as 're-experiencing' or 're-living' a past event. This would seem to imply a phenomenal similarity between memory and perception. But episodic memory is also typically characterized as involving a *feeling of pastness* (Dokic, 2014; Fernández, 2019; Perrin, Michaelian & Sant'Anna, 2020), a feeling that is not present in perception. What, then, is the precise relation between the phenomenal properties of perceiving and remembering?

These questions are among the many important philosophical issues sitting at the intersection between these two areas. The goal of this special issue is to invite a dialogue on the relation between memory and perception, to help shine a light on the relation between these cognitive capacities. The issue is also open to papers that explore the ways in which traditional philosophers have reflected on the relation between memory and perception. In addition, the special issue features a symposium on Jordi Fernández' book *Memory: A Self-Referential Account*. Fernández' monograph provides a functionalist account of memory, which emphasizes and delineates the rich ways in which memory and perception are related. Fernández' book, and the critical engagement it engendered for the symposium, provides a further focal point for shining light on the relation between memory and perception.

2. Contents of the special issue

The papers in the special issue can be grouped into four themes: Guerrero Velázquez and Kirby explore interactions between memory and perception; Andonovski and Sant'Anna examine the vexed issue of whether a causal connection necessarily obtains between remembering and perception; Rosen & Barkasi and Trakas focus on aspects of the phenomenology of memory and how this relates to perception; Díaz Quiroz reflects on memory in the writings of a historical figure. In addition to these submitted papers, we include an invited symposium on *Memory: A Self-Referential Account,* with contributions by Viera, James, and Robins, as well as a response from the author, Fernández. Our symposium is completed with a book review of Fernández' monograph by Álvarez. Here we summarize the contributions.[2]

2.1 Interactions between memory and perception

Memory is in some sense related to perception; the content of perception is somehow retained, restored, or reconstructed in memory, but they are usually considered to be separate and independent cognitive capacities. Two of the papers in the special issue question this complete separation, and look at interactions between memory and perception.

---

2   The order we present the papers here focuses on thematic connections, which is different from the order in which the papers appear in the journal itself.

Carlos Alberto Guerrero Velázquez investigates the relation between memory and perception through the medium of autobiographical interviews. Recognizing the changes that can occur in the content of autobiographical remembering —memory's so-called "creative character"— Guerrero Velázquez suggests that this creativity is related to the simulative nature of episodic memory. Episodic memory is an imaginative capacity for simulating scenarios, which can draw on various sources of information in the process of constructing simulations. This simulative character of episodic memory is revealed, Guerrero Velázquez argues, in autobiographical interviews, where interviewees draw on perceptual information to generate and modify their discourse to respond to a communicative purpose. Memory and perception are not entirely separate and independent capacities, but the processes of remembering can incorporate information from present perception when constructing simulations of the past. The content of memory in this sense is related not only to *past* perceptual content but also to *present* perceptual content.

As an artist using photographic techniques to explore philosophical ideas about memory and identity, **Alun Kirby** adopts a different perspective and explores the inverse interaction between memory and perception —how memory is involved in perceiving aspects of photographs. Drawing on Barthes' notion of the *punctum* in photographs, Kirby shows that a key part of perceiving photographs, and feeling their affective force, is that we view them through the lens of our own autobiographical memories. The punctum of the photograph pierces our memories and releases its affective force. This typically unfolds in an associative way, and Kirby uses this notion of association in his artwork to embody an analogous process to memory. Kirby describes the metamorphogram: a non-traditional photograph, which fits his defined criteria for being analogous to memory. In particular, the metamorphogram, for Kirby, is analogous to a composite of the entirety of an individual's episodic memories. Kirby then uses the insights gained from creating and contemplating metamorphograms, individually and in other artistic works, to consider a bi-directional relationship between individual autobiographical memory and shared cultural memory. Kirby's artistic and philosophical journey into memory and identity leads him to propose a new form of memory, which is embodied in particular groups such as collectives of football fans. Kirby calls the form of memory he has identified *sociobiographical memory*.

2.2 Memory and perception: causal connections?

The question of whether an appropriate causal connection is necessary for remembering is one key issues in philosophy of memory.[3] Indeed, it is sometimes thought that the debate about the relationship between memory and imagination is ultimately a question

---

3   For a nice summary, see Michaelian & Robins (2018).

of whether such a causal connection is necessary for successful remembering (Perrin & Michaelian, 2017). Two papers in the special issue broach this issue of causation in memory.

Nikola Andonovski focuses in on Fernández' (2019) account of episodic memory and how it speaks to the issue of causation (see also section 2.5 below). On Fernández' functionalist theory of memory, a given state need not be caused by a past experience to qualify as a memory. Rather, in order to be classified as a memory, the state needs to play a particular functional role, the role that a memory state plays. On this functionalist account, memories *tend to be* caused by past perceptual experiences, but causation is not a necessary condition. According to Andonovski, this functionalist account does not progress the debate about memory causation. Andonovski argues that for a given mental state to count as realizing a particular functional role, it must be embedded in an appropriate kind of system. In the case of episodic memory, such a system, Andonovski suggests, is one that supports the kinds of interactions that map onto the relations specified by causal theories and denied by simulationist theories. The crux of the issue is that whether or not there exists the type of system that would support the functional characterization of a state as a memory is largely an empirical matter. This empirical question will not be settled by Fernández' functionalist account, and so the theoretical gains of endorsing it would be, according to Andonovski, minimal.

André Sant'Anna also centers his discussion on causality and the (dis)continuism debate. According to discontinuists, memory is a different kind of state or process to imagining; according to continuists, in contrast, there is no fundamental difference in kind but merely differences in degree.[4] This debate currently centers on the necessity of an appropriate causal connection between memory and past perceptual experience, with discontinuists (causalists) affirming, and continuists (simulationists) denying, the necessity of causation in remembering. Sant'Anna offers a new way of thinking about this debate. He develops an argument based on an analogy to perception, according to which representationalism about perception can accept that two states (veridical and non-veridical perceptual experiences) differ in terms of a causal connection *without* this marking a difference in kind. Sant'Anna then proposes that, given the commitment by causalists and simulationists to a representationalist approach to mental states, it is spurious to frame the (dis)continuism debate in terms of the necessity of a causal connection. Instead, Sant'Anna proposes, the debate should be viewed as being about the nature of the *attitudes* involved in remembering and imagining as opposed to their contents. If remembering and imagining involve different attitudes, then discontinuism prevails. If, on the other hand, it can be shown that remembering and imagining involve similar attitudes, then continuism is correct. Sant'Anna's view hence offers a distinct perspective on the (dis)continuism debate.

---

4   See, for example, Michaelian et al. (2020) for a summary of the debate.

2.3 The phenomenology of memory and perception

What is the precise relation between the phenomenal properties of perceiving and remembering? Does the phenomenology of episodic memory depend on, or relate to, in some sense, past perceptual experience? These are the topics of focus for two papers in the special issue.

Melanie Rosen and **Michael Barkasi** describe similarities and differences in the phenomenology of memory and perception. They note that episodic memories typically involve a feeling of pastness, whereas perceptions typically involve a feeling of presence. Rosen and Barkasi show how these phenomenology feelings can come apart in some cases from memory and perception, and that the feelings themselves cannot be used to ontologically ground experiences as memories or perceptions. Despite this, they argue that the feeling of pastness and presence are genuine features of episodic memories and perceptions, respectively, and that they are important characteristic markers that help us categorize mental states first-personally. They provide an account of the feeling of pastness in episodic memory that is distinctly phenomenal, rather than doxastic, although as they show, our web of beliefs may also contribute to phenomenal experience. Indeed, according to Rosen & Barkasi, the feeling of pastness of episodic memory and the feeling of presence of perception are intimately connected. They outline several cognitive features that underlie both phenomenological feelings, including the feeling of (past) accessibility, ergonomic significance, immersion, objectivity, and mental strength. In this way they offer a novel account of what grounds the feeling of pastness in episodic memory and highlights its relation to the feeling of presence in perception.

Marina Trakas also focuses on an aspect of the phenomenology of episodic memory. Trakas' starting question is: What does it take for a subject to experience a memory as being her own? She then critically examines Fernández' (2019) answer to this question. According to Fernández' endorsement account, this particular phenomenal quality of our memories can be explained in terms of the experience of mnemonic content. When the subject feels a memory as her own it is because the memory has been caused by her perception of a past veridical fact. The memory matches past perceptual experience, it is represented as veridical, and hence the memory is endorsed. It is this endorsement of the memory that Fernández thinks results in the feeling of mineness. Trakas thinks that the endorsement model does not have sufficient explanatory value to account for the sense of mineness of our memories. She outlines two major worries with the endorsement model. First, she suggests that the evidence Fernández uses to empirically ground his theoretical proposal is found wanting. Fernández bases his theoretical analysis of the sense of mineness in the linguistic analysis of some reports of the patient R. B. (Klein & Nichols, 2012), who is a person who, for some time after suffering an accident, claimed to have memories but not feel that these memories were his own. Trakas outlines worries about building an account of the sense of mineness

from R. B.'s reports of his memory experiences. Second, Trakas appeals to the empirical literature on non-believed memories (Mazzoni, Scoboria & Harvey, 2010) to suggest that the endorsement model fails to accommodate many non-pathological everyday memories that preserve their sense of mineness, but whose accuracy is explicitly denied, suspected or not automatically endorsed. In these cases, subjects know the memories are not veridical, and hence do not endorse them, but they are memories that seem to maintain a feeling of mineness. According to Trakas, these non-believed memories are counterexamples to the endorsement model of the sense of mineness.

2.4 Reflections on a historical figure

Adopting a historical perspective, **Diego Díaz Quiroz** looks at the notion of intellectual memory in Descartes. Díaz Quiroz notes that Descartes recognized both a bodily form of memory, which can be explained in physiological terms, and an intellectual or spiritual form of memory, which doesn't reside in any bodily organ. Díaz Quiroz's concern is with this latter form of memory, and whether it was used by Descartes for philosophical or theological reasons. Drawing on an analysis of written correspondences between Descartes and Burman, and Descartes and Arnauld, Díaz Quiroz details the ways in which Descartes explains the concept of intellectual memory. He argues that, for Descartes, corporeal memory can register the physical aspects of sense impressions, such as the sounds of words, but not their meaning. It is intellectual memory that helps the mind understand the meanings behind sense impressions. Indeed, Díaz Quiroz goes on to suggest that corporeal memory and intellectual memory operate together to produce processes of reminiscence in human beings. Hence, Descartes' understanding of intellectual memory is primarily epistemic and semantic, and not, Díaz Quiroz concludes, theological.

2.5 Book Symposium – *Memory: A Self-Referential Account*

Jordi Fernández offers an engaging and sophisticated philosophical account of episodic memory. Fernández proposes that memories play a particular functional role in our cognitive economies, one which involves past perceptual experiences and beliefs about the past. Given the importance of the relation between memory and perception, on Fernández' account, as part of this special issue we invited a series of commentaries on Fernández' manuscript, as well as a response from Fernández to this critical engagement with his work.

    Gerardo Viera focuses his discussion on time in memory. Episodic memories often come with a feeling of pastness. The events we remember feel to us as if they occurred in our pasts. This raises the immediate question: what accounts for this

feeling of pastness? Viera's paper aims to provide an answer to this question. Viera first raises objections to Fernández' account of this phenomenological aspect of episodic memory. According to Fernández, the feeling of pastness is not due to an explicit representation of the temporal location of the remembered event. Instead, for Fernández, the feeling of pastness is grounded in the self-referential causal content of memory. That is, memory represents its own causal origin, and it is this that grounds the feeling of pastness. Viera thinks that that this account falls short. He notes that causation and time are closely related but they are not identical. Drawing on empirical evidence showing that, in perception, experiences of causal and temporal order can come apart, Viera argues that the self-referential view fails to explain why a representation of cause would give rise to a feeling of pastness without introducing an explicit representation of time. According to Viera, the feeling of pastness is better explained by a special form of egocentric temporal representation, which he calls *path-dependent representation*.

Steven James' point of engagement with Fernández lies in the nature of judgements based on episodic memory and whether they are immune to certain kinds of errors of misidentification. Immunity to errors of misidentification (IEM) refers to the idea that there are certain kinds of judgments in which it is impossible to be wrong about the identity of the person about whom one is making the judgment (Recanati, 2012). For example, I can be wrong about the person whom I judge is giving a talk that I am listening to, but I cannot be wrong that I am the person whose auditory experience is one of hearing a talk. First-personal judgements, based on information gained from the inside, seem to exhibit the property of IEM. According to Fernández, episodic memory is importantly immune to such errors of misidentification. James takes issue with this claim. He first outlines the view that, on the face of it, memory judgments can be prone to errors of identification. He then shows why, on Fernández' view, these cases don't count against the thesis that episodic memory judgments are IEM, and why this ultimately proves problematic. As James shows, crucially, for Fernández, a failure of IEM requires a misidentification on the basis of grounds that are completely and fully accurate. Because the IEM status of memory-based judgments is grounded in completely accurate memory content, James suggests, then there is a certain circularity to Fernández' account of IEM. In other words, the worry is that Fernández' understanding of episodic memory content *ensures* that judgments based on it are IEM. James then concludes that, while technically immune to error through misidentification, episodic memory judgments are not grounded in such a way that they have any interesting epistemological import for the subject in terms of self-consciousness. Rather, insights about our self-conception are directly derivable from the metaphysics of memory content alone.

Sarah Robins raises challenges for a particular aspect of Fernández' account, as well as outlining a more general problem with his application of functionalism to episodic remembering. First, according to Robins, there are concerns about the

mnemonic role that Fernández designates for mental images. In characterizing the mnemonic role of episodic remembering, Robins argues, Fernández fails to make clear how the mental image type that plays this role should be identified. Robins outlines various options for how this identification may be approached, and illustrates specific challenges for each. Robins' worry, in a nutshell, is that evaluating the mental image's mnemonic role involves assessing two tendencies —what tends to cause the mental image and what the mental image tends to cause. The problem results from a tension in that the understanding of the image that is used to establish one side of its mnemonic role looks different to the understanding of the image that is used to establish the other. In order to play its backward role, the image needs to be generic, but in order for the image to play its forward role it needs to be specific. For Robins, there appears to be no stable conception of the mental image that can play both backwards and forwards roles. Robins then raises a more general worry for Fernández' functionalist approach. Such an approach, she thinks, is ill-suited to the metaphysical question about episodic remembering that is of interest to the causal and narrative theorists with which Fernández engages. For Robins, functionalism characterizes mental state *types* and is suited to assessments of episodic remembering *across* individuals, whereas the concerns of causal and narrative theories involve *token* states of remembering and assessments *within* individuals. Functionalism, for Robins, fails to fit the explanatory demands of episodic remembering.

Jordi Fernández responds to the three comments on his book. He first notes that all three commentators are targeting fundamental assumptions in the account of memory he proposes. The criticisms, if on the mark, would entail proposing a radically different account of memory. For Fernández, such a radical move is not necessary, however. There are, he thinks, possible responses to the objections from all three commentators, which are available within the constraints of the account proposed in his book.

In dealing with James' objection in relation to IEM, Fernández thinks that rather than trivializing the notion, we need the memory to be accurate precisely in order to rule out the possibility that an error occurs just because memory is fallible. For Fernández, it is trivially true that memory is prone to error and that judgments about the identity of objects (including oneself) may go wrong because of the fallibility of memory. The interesting thesis is precisely if memory is IEM or not when it is fully accurate. Moreover, this does not lead to an epistemologically weak thesis, according to Fernández. He does not deny that we may have other grounds for believing that memories have certain contents which involve the self, and which will enable us to obtain knowledge about our self-conception. Fernández accepts this point, but there is, he thinks, a considerable leap, from this point, to the point that the IEM phenomenon in fully accurate memory tells us nothing interesting about our self-conception.

In tackling Viera's concerns about the feeling of pastness in memory, Fernández embraces the distinction raised by Viera (coming from Lewis, 1976) between personal time and external time, but thinks that rather than posing a problem for his account,

this distinction can actually be used to think about the feeling of pastness of memory in causally self-referential terms. The key question, for Fernández' account, is whether the property of a remembered event of being at the causal origin of the relevant memory is the property of being in the past or not. Fernández thinks there may be two different answers to this question for the two aspects of the Lewisian temporal distinction. If we are talking about external time, the answer is no. Fernández then admits that his account becomes an error theory of pastness, albeit one that, he holds, is still explanatory. If we have personal time in mind, however, then the answer is yes. What our (personal) memories represent is not that those events have some position in external time, but the fact that those events are causally related to our memories. And the property of being causally related to our current states, such as our memories, is the property of being in our personal past. For Fernández, the notion of personal time meshes well with the commitments of the self-referential account.

In his reply to Robins, Fernández acknowledges the difference between types and token mental images. He thinks, however, that this does not pose a problem for his account. According to Fernández, if we want to know, for a token mental state, whether it counts as remembering that p, it is legitimate to ask whether the mental state is of a certain type, namely, remembering that p. The type of mental state will provide us with the conditions that the token mental state needs to satisfy. While Robins sees the functional and causal theories of memory as talking past each other, with the former focused on mental state types and the latter focused on mental state tokens, for Fernández both are concerned with token mental states. Fernández sees himself as answering the question of whether the token mental state is a state of remembering by considering whether it belongs to a certain type, a type which requires certain conditions to be satisfied. If the causal theorist sees causation not as primarily being about token memory causation, then they, like the functionalist, will also consider whether the token mental state under consideration belongs to a certain type. The conditions required by causal theories for belonging to that type will be different to those proposed by Fernández —involving, for example, probabilities. For Fernández, depending on the way causal theorists see causation, they may be engaged in the same kind of project as his functionalist theory: both theories will be trying to account for the backward-looking conditions that a token mental state needs to satisfy in order to be classed as an episodic memory.

To round off the book symposium, **Juan Álvarez** provides a review of *Memory: A Self-Referential Account*. Álvarez offers a thorough description of Fernández' view. Álvarez notes that there are three main parts to Fernández' monograph: the first deals with both the metaphysics and the intentionality of episodic memory; the second part investigates certain phenomenological aspects involved in episodic remembering; the third part centers on two important debates in the epistemology of memory. Álvarez notes two main contributions to the philosophy of memory that Fernández provides. First, he suggests that Fernández' functionalist theory enriches

contemporary discussions on the metaphysics of memory, which are dominated by causal and simulation theories. Second, he suggests that Fernández' focus on mnemonic content as a source of theoretical tools to clarify other philosophical issues has interesting explanatory potential. Álvarez finally draws on recent criticism of Fernández' account to offer an illuminating and balanced review of *Memory: A Self-Referential Account.*

3. Concluding remarks

The goal of this special issue is to invite a dialogue on the relation between memory and perception, to help shine a light on the relation between these cognitive capacities. Part of this goal is to encourage additional research, providing further insights at the interface of memory and perception. Here we outline just a few of the myriad potential avenues of investigation.

A central debate in the philosophy of perception has pitted representationalists, who hold that perceiving is fundamentally a matter of representing an event, against relationalists, who hold that perceiving constitutively involves a relation to the perceived event. If representationalism is right, then successful perception is not different in kind from hallucination. If relationalism is right, in contrast, then successful perception is indeed different in kind from hallucination. Relationalism thus leads to *disjunctivism.* The analogous debate in the philosophy of memory has been less active but has recently been picking up steam (Aranyosi, 2020; Debus, 2008; Moran, forthcoming; Sant'Anna, 2020, forthcoming; Schwartz 2018). Is remembering fundamentally a matter of representing an event, or does it constitutively involve a relation to the remembered event? While the representationalism-relationalism debate about memory is to some extent analogous to the representationalism-relationalism debate about perception, it is unclear how a present memory might be constituted in part by an event located in the past, and disjunctivism about memory thus raises issues distinct from those raised by disjunctivism about perception.

Another ongoing debate concerns the relationship to perception that is required for accurate remembering. Bernecker (2010) distinguishes between two forms of accuracy in memory: a memory is *true* if it is accurate with respect to the remembered event, and it is *authentic* if it is accurate with respect to the subject's original experience of the remembered event. A memory is authentic, in other words, if its content matches the content of the corresponding perceptual experience. Truth and authenticity can come apart, and there is, at present, no agreement on whether both forms of accuracy are required for successful remembering. Bernecker himself takes both truth and authenticity to be required, as does McCarroll (2018). These authors hold that successful memory must get things right both with respect to what happened and with respect to the subject's perception of what happened. Michaelian (2020), meanwhile, has argued

that only truth is necessary —i.e., that successful memory need only get things right with respect to what happened. An interesting view, the merits of which have so far not been considered in the literature, is that only authenticity is necessary —i.e., that successful memory need only get things right with respect to the subject's perception of what happened. There is thus room for additional work on the relationship to perception that is required for accurate remembering.

Another, related, topic worthy of further attention is the generativity of episodic memory. There are two main senses in which memory may be generative: psychologically and epistemically. Psychological generativity is when new content that was not available during perception added to the memory representation. Given the empirical evidence on constructive memory, is seems clear that memory can be generative in this sense. It further interesting question relates to epistemic generativity. Can memory generate new knowledge or justification that wasn't available in perception, or does the perceptual justification of beliefs transfer to memory? What, in other words, is the epistemological relation between perception and memory? Indeed, does psychological generativity entail epistemic generativity?

Other suggestions for further research would be to look at memories of traumatic events, especially those associated with posttraumatic stress disorder. Such (pathological) memories of trauma can often present a quite different phenomenology to everyday episodic memories. Trauma memories can sometimes feel as if they are happening in the present moment, rather than the past, and can be accompanied by the feeling of presence typically found in perceptual experience (Ehlers, Hackmann, & Michael, 2004). As such there seems to be a shared phenomenology between memory and perception, which might help further illuminate or substantiate the connection between the two. Further, it might be interesting to explore how perceptual impairments can affect memory. For example, Oliver Sacks discusses the case of a painter who became colourblind after suffering an accident. Interestingly, the painter was not only impaired in his perception of colour but was also impaired in remembering colour. As Sacks explains:

> Color perception had been an essential part not only of Mr. I.'s visual sense, but his aesthetic sense, his sensibility, his creative identity, an essential part of the way he constructed his world —and now color was gone, not only in perception, but in imagination and memory as well (Sacks, 1995, p. 35).

The painter suffered an impairment of colour vision that impacted his memory. Yet, even in benign cases, it might be interesting to consider the way in which the colours of the visual world are the same or different in the remembered world. Coloured images may be more memorable (Wichmann, Sharpe & Gegenfurtner, 2002) but are all of our memories coloured like perceptions? Further research may help shed light on the relation between memory and perception in terms of more specific properties such as colour.

These are just some of the interesting issues at the intersection of memory and perception. We look forward to continuing the conversation about the relation between these two fundamental cognitive capacities.

Christopher Jude McCarroll
Centre for Philosophy of Memory,
Université Grenoble Alpes, France

Kourken Michaelian
Centre for Philosophy of Memory,
Université Grenoble Alpes, France

Santiago Arango-Muñoz
Instituto de Filosofía,
Universidad de Antioquia

References

Addis, D. R. (2018). Are episodic memories special? On the sameness of remembered and imagined event simulation. *Journal of the Royal Society of New Zealand, 48*(2-3), 64–88. https://doi.org/10.1080/03036758.2018.1439071

Aranyosi, I. (2020). Mental time travel and disjunctivism. *Review of Philosophy and Psychology, 11*(2), 367–384. https://doi.org/10.1007/s13164-020-00467-9

Bernecker, S. (2010). *Memory: a philosophical study.* Oxford University Press. https://doi.org/10.1093/obo/9780195396577-0072

Bernecker, S. (2015). Visual memory and the bounds of authenticity. In D. Moyal- Sharrock, V. Munz & A. Coliva (Eds.), *Mind, language, and action: proceedings of the 36th international Wittgenstein Symposium* (pp. 445–463). De Gruyter.

Bernecker, S. & Michaelian, K. (Eds.). (2017). *The Routledge Handbook of philosophy of memory.* Routledge. https://doi.org/10.4324/9781315687315

Casey, E. S. (1991/2004). *Spirit and soul: essays in philosophical psychology.* Spring Publications.

De Brigard, F. (2014). Is memory for remembering? Recollection as a form of episodic hypothetical thinking. *Synthese, 191*(2), 155–185. https://doi.org/10.1007/s11229-013-0247-7

Debus, D. (2007). Perspectives on the past: a study of the spatial perspectival characteristics of recollective memories. *Mind and Language, 22*(2), 173–206. https://doi.org/10.1111/j.1468-0017.2007.00305.x

Debus, D. (2008). Experiencing the past: a relational account of recollective memory. *Dialéctica, 62*(4), 405–432. https://doi.org/10.1111/j.1746-8361.2008.01165.x

Dokic, J. (2014). Feeling the past: a two-tiered account of episodic memory. *Review of Philosophy and Psychology, 5,* 413–426. https://doi.org/10.1007/s13164-014-0183-6

Ehlers, A., Hackmann, A. & Michael, T. (2004). Intrusive re-experiencing in post-traumatic stress disorder: phenomenology, theory, and therapy. *Memory, 12*(4), 403–415. https://doi.org/10.1080/09658210444000025

Fernández, J. (2015). What are the benefits of memory distortion? *Consciousness and Cognition, 33,* 536–547. https://doi.org/10.1016/j.concog.2014.09.019

Fernández, J. (2019). *Memory: a self-referential account.* Oxford University Press. https://doi.org/10.1093/oso/9780190073008.001.0001

Klein, S. B. & Nichols, S. (2012). Memory and the sense of personal identity. *Mind, 121,* 677–702. https://doi.org/10.1093/mind/fzs080

Lewis, D. (1976). The paradoxes of time travel. *American Philosophical Quarterly, 13*(2), 145–152.

Martin, C. B. & Deutscher, M. (1966). Remembering. *Philosophical Review, 75*(2), 161–196. https://doi.org/10.2307/2183082

Mazzoni, G., Scoboria, A. & Harvey, L. (2010). Nonbelieved memories. *Psychological Science, 21*(9), 1334–1340. https://doi.org/10.1177/0956797610379865

McCarroll, C. J. (2018). *Remembering from the outside: personal memory and the perspectival mind.* Oxford University Press. https://doi.org/10.1093/oso/9780190674267.003.0007

Michaelian, K. (2016). *Mental time travel: episodic memory and our knowledge of the personal past.* MIT Press. https://doi.org/10.7551/mitpress/10591.001.0001

Michaelian, K. (2020). Confabulating as unreliable imagining: In defence of the simulationist account of unsuccessful remembering. *Topoi, 39*(1), 133–148. https://doi.org/10.1007/s11245-018-9591-z

Michaelian, K., Perrin, D. & Sant'Anna, A. (2020). Continuities and discontinuities between imagination and memory: the view from philosophy. In A. Abraham (Ed.), *The Cambridge Handbook of imagination* (pp. 293–310). Cambridge University Press. https://doi.org/10.1017/9781108580298.019

Michaelian, K. & Robins, S. K. (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

Moran, A. (Forthcoming). Memory disjunctivism: a causal theory. *Review of Philosophy and Psychology.*

Nanay, B. (2018). Multimodal mental imagery. *Cortex, 105,* 125–134. https://doi.org/10.1016/j.cortex.2017.07.006

Perrin, D. & Michaelian, K. (2017). Memory as mental time travel. In S. Bernecker & K. Michaelian (Eds.), *The Routledge Handbook of philosophy of memory* (pp. 228–239). Routledge. https://doi.org/10.4324/9781315687315-19

Perrin, D., Michaelian, K. & Sant'Anna, A. (2020). The phenomenology of remembering is an epistemic feeling. *Frontiers in Psychology, 11*(1531), 1–14. https://doi.org/10.3389/fpsyg.2020.01531

Recanati, F. (2012). Immunity to error through misidentification: what it is and where it comes from. In S. Prosser & F. Recanati (Eds.), *Immunity to error through misidentification: new essays* (pp. 180–201). Cambridge University Press. https://doi.org/10.1017/CBO9781139043274.011

Robins, S. K. (2016). Misremembering. *Philosophical Psychology, 29*(3), 432–447. https://doi.org/10.1080/09515089.2015.1113245

Sacks, O. (1995). *An anthropologist on mars: seven paradoxical tales.* Vintage.

Sant'Anna, A. (2020). The hybrid contents of memory. *Synthese, 197*(3), 1263–1290. https://doi.org/10.1007/s11229-018-1753-4

Sant'Anna, A. (Forthcoming). Unsuccessful remembering: a challenge for the relational view of memory. *Erkenntnis.* https://doi.org/10.1007/s10670-020-00261-0

Schwartz, A. (2018). Memory and disjunctivism. *Essays in Philosophy, 19*(2), 1-18. https://doi.org/10.7710/1526-0569.1611

Shanton, K. & Goldman, A. (2010). Simulation theory. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*(4), 527–538. https://doi.org/10.1002/wcs.33

Sutton, J. (2010). Observer Perspective and acentred memory: some puzzles about point of view in personal memory. *Philosophical Studies, 148*(1), 27–37. https://doi.org/10.1007/s11098-010-9498-z

Wichmann, F. A., Sharpe, L. T. & Gegenfurtner, K. R. (2002). The contributions of color to recognition memory for natural scenes. *Journal of Experimental Psychology: Learning, Memory and Cognition, 28*(3), 509–520. https://doi.org/10.1037/0278-7393.28.3.509

Praxis Filosófica

Universidad del Valle

Memoria y percepción en la entrevista autobiográfica: una simulación episódica que se adapta en tiempo real al contexto*

Carlos Alberto Guerrero Velázquez

Universidad Autónoma del Estado de Morelos, Morelos, México
guerrerovelazquez@gmail.com

Resumen: Normalmente se piensa en la percepción y la memoria como dos capacidades independientes, creyendo que la primera solo tiene influencia sobre la segunda durante la codificación. En las entrevistas autobiográficas de historia oral y memoria histórica, los entrevistados seleccionan, adaptan y completan sus recuerdos para crear diferentes versiones de ellos. En este artículo se argumenta que lo anterior es consecuencia de la naturaleza simulativa de la memoria episódica, y del empleo por los entrevistados de información perceptiva para generar y adaptar sus recuerdos a un discurso autobiográfico, buscando satisfacer un propósito comunicativo. Para ilustrar esto, se analizan tres factores contextuales que influyen sobre la construcción del recuerdo en una entrevista autobiográfica (objetivo comunicativo, idioma de comunicación e interacción emocional), mostrando que, en este tipo de recuperación, memoria y percepción contribuyen simultáneamente a construir simulaciones de eventos que se adaptan en tiempo real al contexto donde ocurre la evocación.

Palabras clave: entrevista autobiográfica, teoría de la simulación, memoria y percepción, memoria constructiva, recuerdo autobiográfico

---

A R T Í C U L O
D E   I N V E S T I G A C I Ó N

Memory and perception
in autobiographical interviews:
an episodic simulation that adapts
in real-time to the context

Abstract: Perception and memory are usually thought to be two independent facul-ties, where the former is believed to only have an influence on the latter at encoding. In autobiographical interviews of oral history and historical memory, interviewees select, adapt, and complete their memories to create different versions. This paper argues that this process is a consequence of the simulative nature of episodic memory and the interviewees' use of perceptual information to generate and adapt their memories to an autobiographical discourse with the goal of satisfying a communicative purpose. To illustrate this, three contextual factors (communicative aim, language of communi-cation, and emotional interaction) that influence the construction of a memory in an autobiographical interview are analyzed, showing that, in this type of recall, memory and perception simultaneously contribute to constructing episodic simulations of events, which adapt in real-time to the context in which recall happens.

Keywords: autobiographical interview, simulation theory, memory and perception, constructive memory, autobiographical remembering

Carlos Alberto Guerrero Velázquez

es Licenciado en Filosofía (UAA – México), Especialista en Ciencias Sociales con Mención en Historia Social (UNLu – Argentina), y Maestro en Ciencias Cognitivas (UAEM – México). Sus intereses de investigación giran alrededor de los temas de identidad, memoria, len-guaje, mente y cognición, tendiendo siempre al enfoque interdisciplinar y al encuentro entre la investigación empírica y la reflexión filosófica.

ORCID: 0000-0002-1384-7601

Check for
updates

A mi madre Leticia y mi padre Alberto, que son y serán siempre la más valiosa y amada de mis memorias.

Introducción

Los atentados terroristas del 11 de septiembre de 2001 que destruyeron el World Trade Center de Nueva York (conocidos popularmente como 11-S) fueron un capítulo importante de la historia contemporánea y la memoria histórica del pueblo estadounidense. Una parte del suceso pudo ser seguido en tiempo real y a nivel internacional por la cobertura televisiva que se hizo de él. Siete semanas después del siniestro, Pezdek (2003) realizó un estudio en el que participaron estudiantes universitarios, cerca de la mitad de ellos residentes de Manhattan (lugar del ataque). En dicho estudio se hicieron entrevistas autobiográficas (aquellas que tienen por objeto conocer los recuerdos de una persona sobre uno o varios episodios de su vida) con el fin de analizar los recuerdos de los participantes con relación al 11-S. Una de las preguntas realizadas fue: "El 11 de septiembre, ¿vio en la televisión el video del primer avión [secuestrado] golpeando la primera torre?". El 73% de los participantes contestó que sí, a pesar de que eso no era posible, debido a que las filmaciones de la primera nave impactando el WTC no se hicieron públicas sino hasta un día después de los sucesos.

El anterior es un claro ejemplo de uno de los fenómenos más comunes en las entrevistas autobiográficas: al evocar un suceso, las personas comunican sus recuerdos elaborando versiones variadas de ellos a partir de la incorporación u omisión de detalles y características pertenecientes a otros episodios o a información relacionada con el incidente. En las entrevistas también es común que, además de recordar, los entrevistados realicen otros procesos cognitivos como evaluar comportamientos o pensar de modo contrafáctico (Arfuch, 1995; Grele, 1989; Portelli, 2017). Todo lo anterior se conoce en estas disciplinas sociales como "carácter creativo del recuerdo" (Arfuch, 1995; 2007) y pone de relieve el talante al menos parcialmente ficcional del discurso episódico que se genera en las entrevistas autobiográficas. En estas, el proceso de evocar se lleva a cabo en una interacción dialógica (en conversación con el entrevistador) y muchos detalles de la elaboración del recuerdo se hacen públicos, permitiendo a un observador externo hacer un seguimiento del curso de la evocación y de la forma en que se construye en tiempo real el discurso autobiográfico.

El carácter creativo del recuerdo ha sido de gran interés para los teóricos de la entrevista autobiográfica (Arfuch, 1995; Grele, 1989; Robin, 1996), y supone problemas epistemológicos y metodológicos que no han sido aún tratados con suficiencia; como la causa de la diferencia entre episodio vivido y recordado, el papel activo o pasivo de la memoria durante la recuperación, y la influencia de las aferencias interoceptivas y exteroceptivas durante la evocación. Entre los teóricos de la entrevista social, este

Estud.filos n.º 64. Julio-diciembre de 2021 | pp. 21-45 | Universidad de Antioquia | ISSN 0121-3628 | ISSN-e 2256-358X

https://doi.org/10.17533/udea.ef.n64a02

23

fenómeno ha sido explicado principalmente como una consecuencia de la interacción dialógica que se da entre entrevistado y entrevistador, aunque no es común, sin embargo, que historiadores orales e historiadores de memoria histórica (Cohen, 2006; Grele, 1991; Portelli, 2017), profundicen en los procesos psicológicos o neurológicos que subyacen a la evocación y los factores que la afectan —debido principalmente a que su enfoque de estudio y su nivel explicativo es otro. Tales procesos han sido estudiados más bien por psicólogos y filósofos de la memoria (Conway, 2005; 2015; Michaelian, 2016; Michaelian & Sant'Anna, 2019; Schacter, 2003), aunque su análisis se ha enfocado comúnmente en la evocación en general y no en la evocación autobiográfica en un contexto de entrevista. Además, estos análisis, en el caso de los psicólogos, comúnmente se centran en la actividad mental del individuo (más específicamente, intracraneal) y dejan de lado los factores externos o contextuales que puedan estar influyendo en el proceso, debido principalmente, a que el diseño de un paradigma experimental exige seleccionar —de entre todas las variables posibles— aquellas que puedan ser controladas y medidas con mayor precisión. En psicología, por ejemplo, la variación entre episodio vivido y recordado ha sido explicada como un efecto provocado por factores como el paso del tiempo, la atribución errónea de la fuente, los sesgos retrospectivos que se dan al recordar (Shacter, 2003); la necesidad de que nuestros recuerdos sean coherentes con la identidad personal o de que sirvan para explicar una situación actual (Conway, 2005).

En este artículo se analiza y describe, desde una perspectiva cognitiva que integra distintos niveles de explicación, el proceso de evocación que se lleva a cabo en la entrevista autobiográfica, con fundamento en literatura de psicología, neuropsicología y filosofía de la memoria. Se pretende mostrar que el carácter creativo del recuerdo autobiográfico en un contexto de entrevista se produce debido a que, al recordar, la memoria humana emplea contenido episódico y semántico para construir versiones del recuerdo que se pretende evocar (Michaelian, 2016; Michaelian & Santa Anna, 2019), y además incorpora información perceptiva interoceptiva y exteroceptiva para dirigir, enriquecer y retroalimentar esa construcción.

Para lograr tal propósito, primero se abundará en la entrevista autobiográfica para describir lo que los historiadores han señalado como el carácter creativo del recuerdo. Posteriormente se abordará la evocación dirigida que se da en un contexto de entrevista, buscando analizar los procesos cognitivos implicados en ella desde las teorías de simulación mnémica (Michaelian, 2016; Schacter, 2018). A continuación se explicará la influencia que la simulación episódica tiene en la construcción de los recuerdos y su relación con el carácter creativo de la evocación autobiográfica. Finalmente, con el propósito de ilustrar cómo es que la memoria emplea información perceptiva para la construcción de la simulación, se describirán tres factores contextuales que influyen sobre ella en el contexto de una entrevista autobiográfica (objetivo comunicativo, idioma de comunicación e interacción emocional). Se mostrará, con lo anterior, que en este tipo de evocación, la memoria y percepción

trabajan simultáneamente para elaborar versiones de los eventos que se adaptan en tiempo real a las necesidades del contexto.

I - La entrevista autobiográfica y el carácter creativo del recuerdo

La entrevista autobiográfica[1] es el suceso comunicativo que se da a partir de la interacción dialógica entre un entrevistador y un entrevistado, el cual es una persona que evoca recuerdos sobre uno o más sucesos de su vida y los organiza en una narración con sentido, llamada "discurso autobiográfico"[2] (Arfuch, 1995; 2007). Se puede considerar un dispositivo epistémico (siguiendo a Fisher, 2006) en tanto que permite a los participantes recordar conocimiento biográfico —tanto episódico como semántico—, organizarlo y crear nuevos contenidos a partir de él.

En las disciplinas de la historia oral y los estudios de memoria histórica se ha señalado en varias ocasiones el carácter creativo del discurso autobiográfico producido en un contexto de entrevista (Barela & García, 2009; Cohen, 2006; Grele, 1989; Pollack & Heinich, 2006; Portelli, 2017), debido principalmente a dos razones. La primera es que no solamente se narran sucesos evocados, sino que se expresan distintas dimensiones de la identidad de una persona, como valores, significación del pasado, sentido histórico, objetivos personales, propósito comunicativo, entre muchos otros (Arfuch, 1995; De Garay-Arellano, 1999; Régine, 1996); para los fines de este trabajo, se nombra a esta característica la *expresión del yo* [EdY]. La segunda razón es que todos estos factores guían de una u otra forma la organización del discurso, la selectividad de los recuerdos a comunicar y su ordenamiento dentro de una narración coherente, en la cual se mezclan la evocación del episodio con otros procesos cognitivos que la persona realiza al momento de recordar, como evaluar acciones propias o ajenas, considerar comportamientos alternativos de acción (pensamiento contrafactual), reflexionar sobre las consecuencias de la actuación, entre muchos otros; característica que en este texto se denomina *realización de procesos cognitivos alternos* [RPCA]: al parecer, las personas tienden a emplear sus recuerdos de un episodio para, además de recordar, construir sentidos relacionados con el suceso o su actuación en este.

Otro aspecto importante de la evocación autobiográfica que se da en contexto de entrevista (y que será profundizado en la tercera sección de este trabajo) es que un entrevistador la guía. Esto supone una creación dialógica del discurso autobiográfico (Arfuch, 1995; 2007), pues la construcción del recuerdo se da en interacción con el otro y no en una vivencia personal privada, como sucede en la evocación espontánea.

---

1   Pueden verse ejemplos de este tipo de entrevistas y su método en Arfuch (1995; 2007); Actis, Aldini, Gardella, Lewin & Tokar (2006), De Garay-Arellano (1999), Barela, Miguez & García (2009).

2   Llamado también "autobiografía" o "narración de vida" en la historia oral. En este trabajo se define como la expresión pública lingüística del recuerdo autobiográfico que se construye en un contexto de entrevista, ya sea de un solo recuerdo o de varios.

En la entrevista, la evocación y la elaboración del discurso autobiográfico se ven influenciados por la presencia del entrevistador, la idiosincrasia de este, las expectativas que fija sobre el entrevistado, su propia comprensión o su necesidad de mayor detalle. Por ejemplo, en una entrevista que se haga a un integrante de una organización política conocido por su militancia, aunque las preguntas no se hagan directamente con respecto a tal militancia, el aspecto político permeará muchas de las respuestas, las cuales serán interpretadas de cierta manera por las opiniones políticas que tenga el entrevistador; este elaborará o reorientará sus cuestionamientos a partir de tales opiniones, de la claridad que perciba en el relato, del motivo de la entrevista o de los detalles que requiera para enriquecer o comprender el episodio (Greele, 1989). Lo anterior provocará que el entrevistado oriente y acomode su discurso (amplíe, omita, enfatice o reelabore detalles o ideas de su narración) para participar en tal interacción, satisfacer los requisitos que plantea el entrevistador, y comunicar lo que desea.

A pesar del énfasis que se ha puesto en el carácter creativo y al menos parcialmente ficcional del discurso autobiográfico y de las implicaciones de la interacción dialógica en la entrevista, los teóricos de esta (Arfuch, 1995; Robin, 1996), así como de la literatura en general sobre el método de la historia oral o memoria histórica (como puede verse en Barela, Miguez & García, 2009; De Garay-Arellano, 1999; González-Monteagudo, 2010; Greele, 1989) no profundizan en los procesos psicológicos que se dan en este tipo de evocación, ni en por qué el recuerdo autobiográfico tiene un carácter constructivo. Las teorías de simulación mnémica pueden servir como un recurso que permite describir —desde el punto de vista cognitivo— de qué forma se da la evocación autobiográfica en contexto de entrevista, así como señalar las características que la hacen adquirir tal carácter creativo.

II - El recuerdo como simulación episódica

Hablar de evocación autobiográfica exige primero aclarar en qué consiste el proceso cognitivo de evocación. Se empleará el postulado teórico de las teorías de simulación mnémica por considerarlo un recurso metodológico robusto —y cada vez más sustentado en evidencia experimental— para explicar los distintos fenómenos de la memoria,[3] encontrando acogida tanto en filósofos de la memoria (De Brigard, 2014;

---

3   Existen distintos postulados teóricos sobre memoria que pueden emplearse para cumplir con el propósito de este análisis, ya que actualmente existe un consenso generalizado sobre el carácter constructivo de la memoria humana (véase Michaelian & Sant'Anna, 2019 para una revisión de tales teorías). Para este trabajo se ha elegido la teoría de la simulación mnémica por distintas razones. En primera, por ser un postulado que goza de una aceptación multidisciplinaria. Lo anterior, como segunda razón, conlleva que se haya desarrollado abundante argumentación teórica e investigación empírica en torno a ella (véase Schacter, 2018). En tercera, a diferencia de otros postulados teóricos —como la teoría causal de la memoria

Michaelian, 2016; Michaelian & Sant'Anna, 2019; Shanton & Goldman, 2010), como en psicólogos y neuropsicólogos (Barsalou, 2005; Cheng, Werning & Suddendorf, 2016; Schacter, 2018).

Según Michaelian (2016), al evocar algo los seres humanos empleamos experiencias pasadas relacionadas al suceso que deseamos recordar y las combinamos de forma flexible para generar simulaciones que nos permiten imaginar[4] un episodio; el producto de esta simulación es lo que experimentamos como recuerdo, lo que en otras palabras significa que recordar el pasado consiste en imaginarlo. Este postulado se ha visto reforzado por evidencia empírica que muestra que existe un traslape en las redes neuronales que empleamos para recordar un episodio y para imaginar el futuro u otras situaciones probables (Schacter, 2018). Tanto Michaelian (2016) como Schacter (2007; 2018) coinciden en señalar que tales procesos son realizados por un mismo sistema cuya función, en general, es realizar simulaciones episódicas. Para Schacter (2018), el propósito evolutivo de este sistema, más que recordar, es habilitarnos para emplear experiencias pasadas y usarlas en la simulación de futuros probables, permitiéndonos así calcular con antelación las implicaciones y consecuencias de una posible situación futura. Michaelian (2016) se refiere a este como "sistema de construcción episódica" (SCE, en lo sucesivo) y señala que se trata de un sistema de construcción capaz de realizar distintos procesos (imaginar el futuro, planear acciones, imaginar escenas ficticias, entre otros), siendo la evocación solo uno más de ellos.[5]

Para Schacter (2018), la habilidad de simular experiencias futuras y construir representaciones de eventos, puede contribuir a errores de la memoria, lo que significa que las desviaciones, omisiones, falsas atribuciones y otros efectos que se dan en la evocación autobiográfica,[6] pueden ser una consecuencia de las características propias del sistema que posibilita la simulación y de que su propósito adaptativo no se agote solamente en recordar. Pero esta falibilidad de la memoria en realidad significa una ventaja adaptativa, ya que "no necesitamos registrar cada detalle de cada experiencia,

---

(véase Michaelian y Robins, 2018)–, la teoría de la simulación otorga un rol principal a la imaginación en la construcción del recuerdo. Esto resulta de gran importancia si se considera que la imaginación tiene un papel fundamental en fenómenos de la memoria como la existencia de falsos recuerdos, la ocurrencia de procesos cognitivos alternos durante la evocación o la posibilidad de emplear un recuerdo como un dispositivo epistémico (algunos se explicarán a lo largo del texto).

4   Existe un gran debate sobre en qué consiste la imaginación. El uso que se le da en este texto es como sinónimo de imaginería mental, empleando la definición de Moulton & Kosslyn (2009), para quienes la imaginería es un proceso cognitivo que consiste en la simulación mental de una situación; simulación que se emplea para hacer explícitas y accesibles las consecuencias de una acción.

5   Para Michaelian (2016), existe una distinción entre la memoria episódica, que nos habilita para obtener conocimiento sobre episodios experimentados en el pasado, y la autobiográfica, que no se relaciona con un sistema específico de memoria, sino que emerge de la interacción de otras capacidades más fundamentales, incluyendo la memoria episódica y semántica. En este sentido, la evocación autobiográfica emplea el sistema de construcción episódica para imaginar un episodio pasado haciendo uso tanto de conocimiento episódico como semántico, provocando que se manifiesten diferentes dimensiones de la identidad de una persona al imaginar un recuerdo autobiográfico.

6   Shacter (1999; 2002) estudia estos efectos y los describe ampliamente en su obra *The seven sins of memory.*

sino extraer las características centrales, el significado o la esencia de las experiencias pasadas, lo cual es fundamental para funciones tan importantes como nuestra capacidad para categorizar y generalizar" (Shacter, 2018, p. 6). De acuerdo a los teóricos de la simulación, a diferencia de otras especies animales, la memoria humana ha evolucionado para ser capaz de integrar conocimiento de distintas experiencias y combinarlo para simular las situaciones requeridas. Esta característica evita, entre otras cosas, el enorme gasto de energía que supondría registrar cada episodio vivido de forma detallada y después reconstruirlo o representarlo mentalmente con todo detalle. En otras palabras, el que la memoria humana sea constructiva y no reproductiva significa una enorme ventaja adaptativa.

¿De qué forma se construye una simulación episódica? De acuerdo con Michaelian (2016), la construcción que se realiza en el SCE emplea de forma flexible, junto con conocimiento semántico e información contextual, experiencias de distintos eventos que están codificadas en las huellas de memoria.[7] Siguiendo a De Brigard (2014), estas huellas pueden entenderse como propiedades disposicionales de las redes neuronales a provocar ciertas respuestas. Durante la codificación (también llamada "registro" de la memoria), se da el fortalecimiento de conexiones neuronales debido a la co-activación de diferentes regiones del cerebro. Una huella de memoria es la propiedad disposicional que estas regiones tienen a la reactivación, cuando son estimuladas por la señal correcta, en aproximadamente el mismo patrón de activación que experimentaron durante la codificación (De Brigard, 2014). El SCE emplea, entonces, contenido[8] de múltiples experiencias habilitado por estas huellas de forma flexible para construir una versión del episodio que se desea recordar, permitiendo con ello que imaginemos el recuerdo en cuestión.

Las teorías de simulación mnémica —especialmente el concepto de SCE como lo plantea Michaelian (2016) y que se ha delineado en esta sección— permiten analizar la evocación que se da en las entrevistas autobiográficas y explorar sus particularidades, siendo el carácter creativo del recuerdo autobiográfico la más importante en este trabajo. Como se mencionó en la sección anterior, este carácter se ve reflejado en dos características descritas por los teóricos de la entrevista: la realización de procesos cognitivos alternos (RPCA) y la expresión del yo (EdY).

La RPCA es muy común durante la evocación autobiográfica: las personas entrevistadas tienden a emplear sus recuerdos para distintas tareas, como la

---

7   Las huellas empleadas no necesariamente se originan en el evento que se pretende recordar, por lo que no necesariamente existe una conexión causal entre la experiencia particular vivida y el recuerdo.

8   "Contenido" no debe entenderse como algún tipo de objeto mental que se encuentra almacenado en el cerebro, lo que corresponde al entendido tradicional de "contenido semántico", sino a la representación que se construye a partir de la activación de las redes neuronales que constituyen las huellas de memoria. En este caso, el contenido no existe previo a la activación, sino a partir de esta, y es un componente constitutivo del proceso de recordar, no el recuerdo en sí (entendido como contenido almacenado previamente). Como señalan Michaelian & Sant'Anna (2019), la evocación episódica puede verse como un proceso que involucra la transición de un estado sin contenido (*contentless*) a uno con contenido (*contentfull*).

actualización de significados a partir de la reconstrucción del suceso ("ahora que lo pienso, eso que pasó me ayudó a…"), el pensamiento condicional contrafáctico ("¿y qué habría ocurrido si…?"), la auto-corrección de información ("ahora que lo pienso, no pudo haber sido así, porque en ese momento yo estaba…"), la incorporación de detalles externos a la vivencia de la experiencia ("dieron el golpe cuando yo estaba incomunicado…"), entre muchos otros. Estos procesos son distintos entre sí, pero tienen en común el estar basados en la imaginación de un episodio y en que esa imaginación es empleada para realizar distintas tareas que, en un momento dado, nos pueden servir para imaginar escenarios posibles o ensayar consecuencias de la actuación. Se puede considerar que la RPCA que se da durante la evocación autobiográfica es una consecuencia natural del funcionamiento del SCE, pues el contenido[9] de la imaginación episódica, aunque haya sido integrado para recordar, permite (e invita a) que se realicen distintos procesos, combinándose de formas variables para generar predicciones, intenciones, planificación, etc. (Szpunar, Spreng & Schacter, 2014). En otras palabras, los seres humanos realizamos otros procesos cognitivos alternos al momento de recordar debido a dos razones: la primera, que todos esos procesos son habilitados por el mismo sistema; y la segunda, que el propósito evolutivo del mismo, más que recordar, es permitirnos imaginar situaciones posibles, razón por la cual empleamos el contenido evocado para ello.

La segunda característica del carácter creativo del recuerdo, la EdY, se refleja en la expresión de múltiples dimensiones de la identidad de una persona que se da durante el recuerdo autobiográfico, como la enunciación de valores personales y grupales, la significación del pasado, la atribución de sentido histórico al suceso, la proyección o construcción que se hace de sí mismo en el episodio, el rol que juega el sujeto en las acciones, etc. Como se ha visto, para generar una simulación, el SCE combina de forma flexible distintas experiencias pasadas, las cuales pueden ser de muy diversos tipos, dependiendo de lo que se pretende recordar; estas experiencias están codificadas en las huellas de memoria y su activación depende de la señal que actúa como disparador. La ocurrencia de la EdY muestra que esta combinación no se limita solamente a experiencias de tipo episódico, sino que la construcción de la simulación se ve afectada por el conocimiento personal (llamada por Conway [2005] la "base de datos del yo") y que la imaginación del pasado tiende a construirse de forma coherente con la identidad personal.[10] A partir de lo expuesto más arriba y siguiendo la teoría de la simulación, es posible suponer que la ocurrencia de la EdY se debe, por un lado, a que las huellas de memoria activadas durante la evocación no se limitan a experiencias sino también al conocimiento personal; y por otro, a que el recuerdo tiende

---

9   Entiéndase "contenido" como el producto de la imaginación episódica, es decir, lo representado en esa imaginación o aquello que se imagina.

10   Conway (2001; 2005) señala que las evocaciones autobiográficas emplean conocimiento sobre la historia e imagen personal de quien recuerda, privilegiando la información referida a sí mismo para generar el recuerdo.

Estud.filos n.º 64. Julio-diciembre de 2021 | pp. 21-45 | Universidad de Antioquia | ISSN 0121-3628 | ISSN-e 2256-358X

29

https://doi.org/10.17533/udea.ef.n64a02

a construirse de forma coherente con la identidad, reforzando así la imagen propia y evitando la disonancia cognitiva (Conway, 2005). Es debido a ello que en los discursos autobiográficos realizados en contexto de entrevista, junto con el recuerdo, a menudo se expresan valores personales ("lo hice a pesar de saber que no era correcto"); valores grupales ("no podía dejar atrás a los compañeros"); asignación de un rol en la historia ("yo siempre fui el tímido de la familia"); imagen de sí mismo ante el suceso ("ahora veo que yo era muy vulnerable entonces"); justificación de las acciones ("me atreví a matar porque era necesario, pero yo no soy un asesino"); entre muchas otras.

Un ejemplo de lo anterior puede verse en la síntesis evaluativa y la proyección de la imagen personal que se hace presente en el discurso autobiográfico de una mujer superviviente de detención ilegal y tortura, que fue recluida en el centro clandestino de la ESMA (Escuela Superior de la Armada) durante la última dictadura argentina. Durante una entrevista autobiográfica donde se le preguntó por momentos de su vida en el centro, mientras narraba episodios de su experiencia, expresó: "Ahora me doy cuenta de que yo, estando adentro, sentía como si me hubieran puesto un vidrio que me separaba del mundo. Sabía que mi nombre no tenía el mismo valor que antes, era un nombre desaparecido" (Gardella, como se cita en Actis *et al.*, 2006). En otros casos se integran conocimientos históricos, sociales o políticos al discurso autobiográfico para complementar y dotar de sentido al recuerdo, pero también para brindar una referencia histórica que ayude a entender el pasado y la forma en que la persona se ve afectada por él. Por ejemplo, en una entrevista realizada a la hija de un militante desaparecido durante la última dictadura argentina, cuando se le preguntó por un recuerdo de sus padres, ella señaló: "Mis padres tenían una vida militante importante. Yo soy de San Salvador de Jujuy y en el año 74 después de que muere Perón, se la llevan detenida a mi mamá, en Jujuy. Ahí deciden que veníamos a Buenos Aires" (Arroyo, citada por Cueto, 2008).

Un último aspecto a considerar en esta parte es la integración de conocimiento semántico y episódico de múltiples recuerdos que se da a menudo en la evocación autobiográfica, y que los historiadores ven reflejada en la construcción del discurso en contexto de entrevista. En el estudio citado al inicio de este texto (Pezdeck, 2003), una gran parte de los entrevistados sobre el atentado del 11 de septiembre al WTC recordó haber visto la imagen del primer avión secuestrado impactando una de las torres, durante la transmisión en vivo del siniestro. Resulta evidente que las imágenes transmitidas televisivamente en días posteriores, a pesar de que no correspondían temporalmente a lo observado el día del siniestro, fueron evocadas también como parte del recuerdo porque forman parte del conocimiento del episodio. Lo mismo sucede en otros casos, donde las personas completan sus recuerdos con información sobre el suceso, detalles de los espacios donde este sucede, y muchos otros datos aprendidos antes o después del episodio. Lo anterior puede obedecer a dos razones. La primera es la forma en que funciona el SCE, y que se ha descrito con anterioridad: al imaginar el episodio, se emplean experiencias pasadas de manera

flexible para generar la simulación, lo cual supone que el episodio imaginado a menudo contenga detalles que podrían corresponder a otros episodios o a información sobre el suceso. La segunda razón es que el propósito adaptativo de la memoria humana es permitirnos imaginar escenarios posibles y calcular las consecuencias de acciones futuras (Schacter, 2018): en términos de supervivencia, es más importante ser capaces de emplear la información más saliente de múltiples experiencias y generar así escenarios de actuación más detallados, que recordar de forma puntual cada detalle de una sola experiencia. En otras palabras, las personas integramos conocimiento de distintas experiencias al recordar porque eso no permite construir simulaciones más detalladas, lo que nos vuelve mejores para predecir.[11]

Esto último abre también la puerta para pensar en otra característica del recuerdo episódico, que es el empleo de información contextual en la construcción de la simulación. La idea de Schacter (2018) de que el propósito de nuestra memoria episódica es imaginar situaciones posibles para predecir consecuencias, conlleva pensar en una memoria motivada (guiada por objetivos) cuyos procesos se llevan a cabo para ayudarnos a mejorar nuestras posibilidades de actuación y éxito en el presente y futuro. Son los estímulos actuales, tanto contextuales como reflexivos, los que activan, motivan y guían nuestras imaginaciones episódicas. En este sentido, resulta lógico pensar que el contenido empleado para la construcción de una simulación no se limita solamente a nuestras experiencias pasadas, sino que incluye también información perceptiva que le permite al agente obtener información de su contexto. Se profundizará en ello en la tercera parte de este artículo.

III - Factores contextuales que influyen en la construcción del recuerdo

Como se mencionó anteriormente, existen distintos factores que influyen en la evocación autobiográfica. En esta parte se tratan aquellos que surgen debido a la situación contextual de la persona que evoca y la interacción comunicativa dinámica que se da entre los hablantes (entrevistador y entrevistado[s]). [12]

---

11 La habilidad de la memoria humana para hacer abstracciones y generalizaciones en lugar de registros precisos puede verse también en otros tipos de memoria. Experimentos que emplean el paradigma Deese–Roediger–Mcdermott (DRM) (Deese, 1959; Roediger & Mcdermott, 1995) muestran que si se presenta una lista de palabras semánticamente relacionadas a una persona ("gato, oso, león, jirafa, mamíferos"), esta tenderá a recordar falsamente una palabra señuelo no incluida en la lista ("animales"). Lo anterior es un ejemplo de que, más que recordar información puntual o detalles específicos, nuestra memoria —en su funcionamiento normal— es mucho más hábil en recordar de una forma temática y global. Esto puede significar una gran ventaja adaptativa en términos de reducción del gasto energético empleado en el proceso cognitivo del recuerdo y de optimización para la construcción de escenarios probables, si seguimos a Schacter (2018).

12 Aunque el propósito de este texto no es abundar en la discusión sobre la naturaleza del lenguaje, es importante recordar la distinción que existe entre el mismo y la comunicación, ya que en distintos animales no humanos se observan sistemas de

Durante la evocación autobiográfica en una entrevista la simulación mnémica se construye y adapta para satisfacer un propósito comunicativo, por lo que el entrevistado incorpora información perceptual contextual que influye en tiempo real sobre la imaginación del episodio. En una entrevista, el entrevistado percibe, decodifica e interpreta estímulos perceptivos visuales, auditivos y táctiles producidos por el entrevistador u otras personas, y se vale de ellos para fijar el objetivo de la indagación mnémica (lo que se le pide que recuerde), conocer las condiciones y detalles que debe tener el reporte del recuerdo, o recibir retroalimentación sobre la calidad de su relato. Esta información le permite también adaptar su construcción episódica "sobre la marcha", elaborando versiones del episodio evocado que se modifican dependiendo de la recepción de su discurso. Los factores que se exponen a continuación muestran que, en la evocación autobiográfica producida durante una entrevista, la memoria se vale de la información perceptiva del contexto para realizar construcciones episódicas que satisfagan las necesidades de actuación del agente.

En la entrevista autobiográfica es especialmente notoria la relación que existe entre percepción y memoria durante la construcción del recuerdo episódico, ya que este tipo de evocación requiere necesariamente la interacción dinámica del agente con su contexto y con los sujetos que forman parte de él. Tal vez podría discutirse si los estímulos perceptivos tienen influencia en la evocación espontánea privada,[13] pero en una situación conversacional (como lo es la entrevista) esta relación se vuelve mucho más evidente porque la interacción con el contexto forma parte constitutiva del proceso de recordar: el estímulo que detona la evocación es externo (pues la pregunta que fija el objetivo de la evocación proviene del entrevistador); la construcción del recuerdo se hace de forma lingüística (empleando el lenguaje oral para narrar, describir, explicar o detallar el evento) y se constriñe a las particularidades del idioma empleado por los interlocutores o las formas de expresión habituales (como las estructura narrativa o el uso de jergas); el entrevistado interpreta el lenguaje verbal y corporal del entrevistador y/o de otros entrevistados porque sabe que su relato está siendo evaluado, etc. Estas características, entre muchas otras, nos muestran la dependencia que tiene la memoria de la percepción para la consecución exitosa del(los) propósito(s) del agente.[14]

---

comunicación sofisticados que, sin embargo, no cumplen con ciertas características que sí tienen los lenguajes humanos. Esta distinción es imprescindible para comprender que aunque el lenguaje es una forma de comunicación, no toda forma de comunicación es un lenguaje (véase Smith & Kosslyn, 2008). En este texto se guarda esta distinción empleando el término "comunicación" como referencia a un intercambio de información (puede ser no verbal), y "lenguaje" cuando se habla de un sistema de comunicación humano que cuenta con un sistema de signos y tiene características específicas (dualidad de patrón, arbitrariedad, capacidad generativa, recurrencia).

13   Aunque si se comparte una visión corporizada, situada o extendida de la cognición, es muy poco probable que se discuta este punto (véase Barsalou, 2008; Fuchs, 2012; Rietveld, Denys & Van Westen, 2018).

14   Algunos ejemplos de los objetivos que se presentan en la entrevista pueden ser: comunicar, convencer, despertar emociones, defender su actuación, congraciarse, verse aceptado, mostrarse con un rol en la historia, mover a la acción, entre muchos otros.

Los factores que se exponen en esta sección se detectaron al analizar las características de la evocación autobiográfica en entrevista, tanto las mencionadas en el párrafo anterior como otras. Los mismos se exponen con el propósito de ilustrar la forma en que la construcción episódica es llevada a cabo por el SCE y cómo se ve influenciada por información perceptiva tanto interoceptiva como exteroceptiva. Tales factores, aunque han sido explorados en distintas investigaciones (las cuales se mencionan a lo largo de esta sección), no se han enfocado en la habilidad específica de la evocación autobiográfica. Debido a las consideraciones de espacio, en este artículo se describen solamente los 3 más representativos del efecto mencionado (objetivo comunicativo, idioma de comunicación e interacción emocional), en lugar de los 6 que se analizaron en la investigación original que sirvió como base para la elaboración de este texto,[15] en espera de abordar los restantes en otra oportunidad.

A) Objetivo comunicativo:

La simulación episódica se construye a partir del contenido habilitado por las huellas de memoria. La activación y selección de tal contenido es guiada por el objetivo comunicativo, que se establece mediante la interacción dialógica.

Como ya se ha descrito, el estímulo que desata la construcción del recuerdo en un contexto de entrevista es la pregunta del entrevistador, la cual es interpretada por el entrevistado para obtener información acerca de lo que se espera de su reporte mnémico. Esta pregunta provoca que la activación de huellas de memoria relacionadas con el suceso (Michaelian & Sant'Anna, 2019) sea selectiva y esté orientada por un objetivo fijado lingüísticamente, a diferencia de la evocación espontánea. Supongamos que un entrevistador pregunta: "¿cuál fue tu reacción al darte cuenta de que el edificio podía caerse durante el sismo?". Esta pregunta ocasionará que se active en el entrevistado conocimiento relativo al suceso, como recuerdos episódicos sobre el sismo en cuestión, información del suceso leída o escuchada en medios de comunicación, conocimiento sobre las reacciones posibles de una persona en situaciones de riesgo, posibles consecuencias de un derrumbe, información sobre su imagen personal (si se considera a sí mismo valiente, arriesgado, temerario, etc.), entre muchos otros. También se activará conocimiento que no es relevante para la indagación y que no será empleado, información que no se desea comunicar, detalles que se descartarán por considerarlos nimios, etc.

---

15   La mencionada investigación fue un trabajo teórico de tesis de maestría en ciencias cognitivas, desarrollado a lo largo de 18 meses. Los seis factores que se analizaron se dividieron en dos secciones, dependiendo de su relación con otras habilidades cognitivas además de la percepción. Un primer grupo lo constituyeron los factores relacionados con el lenguaje: objetivo comunicativo, idioma de interacción y estructura narrativa. El segundo grupo lo formaron aquellos relacionados con las emociones: estado de ánimo, interacción emocional y presión social.

A pesar de disponer de una gran cantidad de contenido, no todo es empleado para imaginar el recuerdo, pues no resulta relevante para el objetivo que se ha planteado. El acto de recordar, en este caso, se asemeja a una búsqueda: tratamos de encontrar los elementos que mejor nos ayuden a construir la escena, conscientes de que nuestro interlocutor espera obtener información de nuestro recuerdo. La interacción lingüística permite, así, que el entrevistador comunique sus expectativas sobre el tipo de recuerdo esperado y la forma en que el discurso debe orientarse; con ello define características del recuerdo como el papel de la persona entrevistada en el suceso que se está reconstruyendo (si este se proyecta como líder, agente de los hechos, observador de los mismos o algún otro personaje dentro de la narración); el objetivo de la indagación (le indica al entrevistado qué se está buscando en el recuerdo); o las características narrativas del recuerdo (si se debe dar una descripción, conducir a una evaluación del suceso, hacer una introspección, etc.). Esta expectativa permite que, a partir del abundante contenido habilitado por la activación de las huellas de memoria, la persona seleccione los elementos que deben formar parte de su narración y que construya la simulación episódica del acontecimiento a partir de esos elementos.

Pero el objetivo comunicativo no solamente se fija al principio de la indagación, sino que se adapta dinámicamente a las necesidades comunicativas que se actualizan, conforme la evocación avanza, mediante la interacción entre los hablantes. Mientras elabora su discurso autobiográfico, el entrevistado recibe y procesa información perceptual, en este caso lingüística, que le permite evaluar el éxito de su comunicación (definido por el nivel de logro planteado) y confirmar, refinar o modificar su objetivo (el motivo que guía la búsqueda) y su estrategia comunicativa (si hay que ampliar, ocultar, enfatizar algún punto, detallar o modificar el relato). Lo anterior ocasiona que el sujeto incorpore en su construcción episódica el contenido necesario para lograr su intención, imaginando una versión del recuerdo que pueda satisfacer la expectativa. Es por ello que los entrevistados pueden hacer un discurso autobiográfico distinto de un mismo suceso si el entrevistador, por ejemplo, expresa que el fin de la entrevista es conocer a las víctimas de un sismo o a los héroes del mismo, ya que el contenido que una persona evoca en situación de entrevista y la forma en que lo hace están mediados por lo que se espera de ella (deseabilidad social); como han reportado los teóricos de la entrevista autobiográfica[16] (Arfuch, 1995; Grele, 1989).

No es de extrañar que lo anterior suceda si tomamos en cuenta las características del ya mencionado SCE (Michaelian, 2019), especialmente el hecho de que su función principal no sea recordar, sino permitirnos generar simulaciones de situaciones posibles

---

16   Lo anterior concuerda también con lo señalado por Conway (2005), quien sostiene que la evocación de un recuerdo está mediada por el propósito que la persona tiene al momento de recordar.

para calcular las consecuencias de una acción (Schacter, 2018). La simulación episódica que se genera en la evocación autobiográfica se orienta por un objetivo, de la misma forma en que una simulación prospectiva o de otro tipo lo harían, pues esa es la función de la simulación en general. Se recuerda *para algo*, con un propósito; en este caso, a saber, para expresar un significado a través del lenguaje.

No se debe suponer, sin embargo, que el objetivo comunicativo de la persona que evoca necesariamente coincida en todo momento con el del entrevistador. La entrevista, como espacio dialógico, está sujeta a una tensión discursiva donde se debaten dos o más subjetividades (Arfuch, 1995; 2007; Grele, 1989). Los entrevistados a menudo toman rumbos evocativos distintos a los programados y planteados por el entrevistador, construyendo su discurso a partir de lo que consideran relevante para ellos del suceso. Estas "digresiones" del objetivo planteado se dan por muy distintas razones: porque ayuda a los entrevistados a significar un episodio; porque notaron un detalle saliente del que desean abundar; porque necesitan justificar sus acciones del pasado; porque quieren expresar una valoración, ampliación o apología del suceso; porque desean abundar en información que les ayude a construir el recuerdo; porque requieren detenerse a argumentar un punto polémico, una opinión sobre el pasado o hasta una defensa de su testimonio; etc.

B) Idioma de comunicación:

La activación de las huellas de memoria parece estar relacionada con el idioma en que se realiza la entrevista. La generación y expresión del recuerdo, al elaborarse de forma lingüística, son influenciadas también por este factor.

El lenguaje oral es quizás uno de los factores más importantes que influyen sobre la evocación autobiográfica en la entrevista, no solamente por la interacción dialógica en que esta sucede, sino porque la construcción misma del recuerdo se realiza haciendo uso de él.[17] Una de las características susceptibles de analizar en la entrevista autobiográfica es el idioma en el que se da la interacción dialógica y la forma en que este podría influir en el proceso de evocación. Dado que una parte importante de las entrevistas autobiográficas del siglo XX (cuando esta metodología cobró auge) se hicieron a inmigrantes, refugiados o grupos indígenas minoritarios, la cuestión del idioma ha sido importante para la entrevista autobiográfica como metodología de la historia oral y la memoria histórica. Para esta última ha tenido especial relevancia, dado que su nacimiento y consolidación como disciplina de las ciencias sociales se dio a partir de la recuperación de discursos autobiográficos de expatriados sobrevivientes

---

17   Determinar relación entre lenguaje y pensamiento, o más específicamente la influencia de uno sobre otro, es una discusión muy abundante (puede revisarse Chomsky, 2006; Gomila, 2012; Smith & Kosslyn, 2008; Whorf, 1956), y ciertamente, no es motivo de este artículo.

de las guerras mundiales, entrevistas que, muchas veces, se han realizado en un idioma distinto a la lengua materna de los entrevistados.

Distintos estudios con personas bilingües (Conway, 2003; Javier, Barroso & Muñoz, 1993; Marian y Neisser, 2000) han determinado que el idioma influye fuertemente sobre la memoria autobiográfica, mostrando que existe un cambio en los patrones de evocación dependiendo del idioma que las personas emplean para elaborar un recuerdo; es decir, que el discurso autobiográfico y los recuerdos evocados son distintos cuando las personas bilingües evocan empleando cada una de las lenguas que hablan. También se ha detectado que las personas interpretan y codifican de distinta forma una experiencia dependiendo del idioma en que están interactuando (Boroditsky, Ham & Ramscar, 2002). Esta influencia parece estar presente incluso en la forma en que se expresan identitariamente las personas bilingües que tienen una doble procedencia cultural, pues construyen diferentes discursos de su yo, dependiendo del idioma que estén empleando para hacerlo (Ross, Xun & Wilson, 1992).

Lo anterior sugiere que nuestros recuerdos autobiográficos están fuertemente ligados al lenguaje (en general) y al idioma (en particular)[18] que empleamos para recordar una experiencia, lo que nos permite suponer que la activación de las huellas de memoria es susceptible a esta influencia. Al parecer, el idioma podría influir en el tipo de recuerdos que son activados para crear la simulación mnémica, estableciéndose una distinción dependiente del idioma al que un recuerdo está asociado. Por ejemplo, en la investigación aplicada con bilingües (Guttfreund, 1990; Marcos, 1976), se encontró que la psicoterapia puede ser más efectiva cuando la lengua de evocación coincide con aquel en que se vivió la experiencia, o que es posible evocar un mayor número de recuerdos cuando una persona lo hace en su primer idioma (Mortensen, Berntsen & Bohn, 2015). Se intuye que, dado que la activación de las huellas de memoria habilita el contenido con el que se genera la simulación, el idioma en el que se está interactuando supone la activación de patrones que entre dos idiomas pueden ser similares o concomitantes, pero distintos, lo que se ve reflejado en la construcción del recuerdo.

El efecto anterior permite analizar también la relación tan estrecha que existe entre el idioma y las emociones. Estudios con bilingües, como los anteriormente citados, muestran que existe una diferencia significativa en la intensidad de las emociones experimentadas por los sujetos cuando interactúan en uno u otro idioma, al realizar tareas que implican el uso de memoria autobiográfica (Guttfreund, 1990; Javier, Barroso & Muñoz, 1993; Marcos, 1976). Es altamente probable que esta diferencia influya también en la construcción del recuerdo, pues la interacción en cierta lengua permite a los sujetos tener una vivencia emocional más intensa, enriqueciendo sensorialmente

---

18   Es importante puntualizar la distinción entre lenguaje e idioma en este trabajo, entendiendo el primero como la facultad comunicativa humana y al segundo como lengua de un grupo humano.

la simulación episódica e incrementando la sensación de re experimentación del evento o viaje mental en el tiempo (Lolich & Azzollini, 2017; Tulving, 2002). Como podemos ver, a diferencia del objetivo comunicativo, el idioma no se relaciona directamente con los estímulos exteroceptivos que influencian la evocación, sino que influye sobre la información interoceptiva que se produce a partir de la experiencia emocional. Los estímulos de este tipo que el sujeto percibe, es decir, la información perceptual interoceptiva que influye sobre la construcción del recuerdo, tiene una intensidad (*arousal*) –y posiblemente una valencia– diferente en cada lengua de interacción.

Esta dependencia del idioma, por lo visto, se daría no solamente en la activación selectiva de cierto contenido en las huellas de memoria, sino también en la generación de la simulación misma, que se realiza con el contenido activado por ellas. Se ha encontrado que las personas hacen una organización lingüística distinta de sus recuerdos autobiográficos dependiendo del idioma en que se estén expresando (Javier, Barroso & Muñoz, 1993), variando la cantidad de detalles, experiencias relacionadas y tipo de recuerdos evocados en cada caso. Al parecer, el idioma influye fuertemente sobre el tipo de construcción del recuerdo que se hace al permitir el uso de distintos recursos lingüísticos, dependiendo de la lengua en que se esté interactuando y del dominio que se tenga de ella: estructuras gramaticales, percepción del tiempo o el espacio, empleo del género gramatical, riqueza léxica, uso de jerga, entre muchos otros.

C) Interacción emocional:

Existe una interacción emocional del entrevistado con el entrevistador, la cual permite que se lleve a cabo un flujo de expresión y retroalimentación no verbal que orienta el discurso.

Una de las características de la evocación en contexto de entrevista, como ya se ha mencionado, es que la misma se realiza en interrelación con otro, lo que supone que la elaboración y expresión del discurso autobiográfico obtiene constantemente retroalimentación por parte del entrevistador. Este último pide a menudo aclaraciones, ampliaciones, detalles o conclusiones sobre el suceso evocado; afirma con palabras, gestos o movimientos cuando comprende lo que se está narrando; reorienta la dirección lógica y afectiva de la evocación (Ortí, 1993), convirtiendo el discurso en el producto de una actividad conjunta (una "negociación", como le llama De Garay-Arellano, 1999). Todo ello provoca una respuesta en el entrevistado, quien adapta su narración como respuesta ante tal comportamiento.

La imaginación del recuerdo está influenciada por la interacción social, pero esa interacción no solamente se da a través del lenguaje verbal, sino también de las expresiones emocionales de los participantes en la conversación, las cuales constituyen otra forma de comunicación. En este caso, la percepción opera de forma similar a como sucede con el primero de los factores mencionados: a través de la captación de estímulos

exteroceptivos que el sujeto decodifica para modular su actuación en el contexto. Estos estímulos le proporcionan a la persona que recuerda información no verbal[19] sobre la forma en que el entrevistador, los escuchas u otros participantes reciben su relato y las emociones que les causa. En la literatura de la metodología de la entrevista autobiográfica en historia oral (De Garay-Arellano, 1999; González-Monteagudo, 2010; Hinojosa, 2013), a menudo se menciona la importancia del *rapport* o la sintonía emocional que debe darse entre los interlocutores para lograr un "contrato de confianza" (Arfuch, 2007) y un contexto seguro que permita el libre flujo de los recuerdos. Este es considerado uno de los pilares técnicos de la entrevista (Hinojosa Luján, 2013) y debe ser propiciado por el entrevistador, quien tiene que procurar fomentarlo y sostenerlo a lo largo de la dinámica. Más que verbal, el *rapport* es una comunicación a través de los gestos, actitudes, posturas, entonaciones y otros elementos de la comunicación que obedecen a una interpretación pragmática. Podemos caracterizarlo como un diálogo que apela a nuestra capacidad para leer las emociones en el otro, sentirnos en confianza y actuar en consecuencia.

La literatura científica sobre nuestra percepción emocional y capacidad de interpretar las emociones de los otros es muy amplia[20] (y su revisión exhaustiva no es motivo de este trabajo). En cualquier caso, estas investigaciones han revelado que constantemente percibimos e interpretamos emociones ajenas (aunque no siempre de forma correcta). Tal percepción se realiza de forma tan sofisticada que no solo somos capaces de interpretar una emoción a partir de los gestos del rostro de una persona, sino también de incorporar detalles del contexto para realizar una correcta lectura sobre la emoción que alguien experimenta o expresa, como rasgos de las escenas visuales, tipos de voz, orientación cultural, entre otros (Feldman, Mesquita & Gendron, 2011). Esta habilidad nos ayuda a gestionar las interacciones con otros, al permitirnos planear, modular y evaluar nuestra acción en contextos sociales. Al igual que con el idioma hablado, la interpretación de las emociones del interlocutor proporciona al entrevistado información que retroalimenta su tarea, en este caso, imaginar su recuerdo. La "lectura" de emociones funciona, entonces, como comunicación no verbal que invita a reforzar, resaltar, aclarar, enfatizar, corregir o reorientar el discurso.

Conclusiones y consideraciones finales

El carácter creativo del recuerdo, que se presenta de forma tan palpable en las evocaciones autobiográficas, ha sido un fenómeno muy discutido por exhibir una de

---

19   Existe controversia sobre si la percepción emocional se realiza de forma separada (modular) de procesos conceptuales como el lenguaje, aunque recientemente se ha sugerido también que el lenguaje puede servir precisamente como un contexto para la percepción emocional (Feldman, Lindquist & Gendron, 2007).

20   Como ejemplos pueden revisarse Atkinson & Adolphs (2011); Keltner & Ekman (2000); Smith & Kosslyn (2008); Tsao & Livingstone (2008).

las características más importantes, pero también más problemáticas de la memoria humana: su carácter constructivo. Esta capacidad de incorporar elementos novedosos a los recuerdos al momento de evocar ha sido considerada en muchas ocasiones un fallo, debido principalmente a que distintos pensadores y científicos de la historia han empleado, para describir los procesos mnémicos, metáforas que comparan la memoria humana con dispositivos de registro, grabación o almacenamiento, caracterizando el recuerdo como una recuperación de información almacenada;[21] lo que se ve reforzado por la concepción popular de la memoria que se da en el lenguaje ordinario. Las teorías de simulación ayudan a entender de otra forma esa capacidad, permitiéndonos ver tal carácter creativo como una característica inherente a nuestra gran capacidad imaginativa, y resaltando la enorme ventaja evolutiva que supone ser capaces de realizar construcciones episódicas con información habilitada por distintos sistemas de memoria a partir de la experiencia vivida.

Como se puede ver a partir de los factores mencionados en este texto, la memoria episódica —al menos durante el proceso de evocación— está fuertemente influenciada por estímulos aferentes interoceptivos y exteroceptivos, lo que muestra que la relación entre memoria y percepción no se limita solamente al momento que llamamos *codificación*. Lo anterior nos lleva a cuestionar qué tanta influencia tiene una sobre otra durante otros procesos cognitivos. Si se toma en cuenta, por ejemplo, que el acto mismo de recordar un evento pone a las huellas de memoria asociadas a este en un estado lábil, como indican los estudios neurológicos de reconsolidación neuronal (ver Labrador & Restrepo-Castro, 2015; Nader & Hardt, 2009; Schwabe & Wolf, 2009), y se considera también que durante la evocación (al menos la episódica) la imaginación emplea información perceptual para recrear el episodio, sería factible sugerir que probablemente gran parte de los recuerdos han sido influenciados un número indeterminado de ocasiones por las situaciones que se viven al momento de recordarlos. Es importante pensar, en futuras reflexiones, las posibilidades y consecuencias epistemológicas de que esta influencia.

Por otra parte, analizar la interacción que se lleva a cabo entre memoria y percepción al momento de la evocación permite enfatizar varias consideraciones importantes.

---

21   Un recuento histórico de las distintas concepciones de la memoria en filosofía —desde la tabla de cera platónica hasta la imagen del pasado wittgensteniana— puede verse en la sección "History of philosophy of memory" del libro *The Routledge Hadbook of Philosophy of Memory* (Bernecker & Michaelian, 2017). En la psicología, aunque las investigaciones de Bartlett (1995) apuntaron tempranamente al concepto de memoria como una habilidad constructiva, el paradigma dominante en los estudios cognitivos durante gran parte del siglo XX empleó la metáfora de la computadora para explicar la cognición humana, caracterizando la memoria como un dispositivo de almacenamiento de información y al recuerdo como un acceso para la recuperación de esta (Baddeley, Eysenck & Anderson, 2020). Aún hoy, que el paradigma computacional ha dado paso a otros enfoques, la terminología dominante para referirse a los procesos mnémicos en esta disciplina continúa empleando palabras ("codificación", "almacenamiento", "recuperación") que reflejan la herencia de tal metáfora (véase Baddeley, 1990; Smith & Kosslyn, 2008).

Primera, la permeabilidad que existe entre procesos cognitivos,[22] lo cuales no se llevan a cabo de forma aislada, sino que son influenciados unos por otros en una interacción dinámica que permite en cada momento la adaptación del agente al contexto. En segunda, muestra que procesos cognitivos tradicionalmente considerados *online* como la percepción y *offline* como la memoria (quizás no haya otro más representativo), en realidad operan de forma simultánea. Un último punto para discutir en futuros trabajos es la relación entre la capacidad humana de efectuar simulaciones episódicas y la habilidad para estructurar de forma narrativa los recuerdos: ¿es constitutiva o causal? Si es esta última, ¿cuál de ellas causa la otra? ¿Qué relación existe entre el desarrollo de los circuitos neuronales que son reclutados para el funcionamiento del SME y el desarrollo como seres narrativos?

Como se puede ver, existen algunos aspectos que han dominado el curso de la argumentación y exploración en este texto. El primero de ellos, es la defensa de la memoria humana como una habilidad cuyo propósito adaptativo y práctico no es el almacenamiento de información. En las entrevistas autobiográficas tradicionalmente se ha entendido a la persona entrevistada como una fuente de información episódica y al entrevistador como un investigador que busca el recuerdo en esa fuente[23] (Aron-Schnapper y Hanet, 1993). Pero la persona que evoca es un agente que tiene un propósito en la interacción dialógica (comunicar su recuerdo, el significado de este, etc.), lo que refuerza la perspectiva de que el papel de la memoria humana no se limita a volver disponible la información del pasado, sino a permitirnos actuar en el mundo haciendo uso de esa información (Schacter, 2018). Recordar, desde esta perspectiva, obedece a una necesidad de actuar para cumplir un propósito. La profundización en este rasgo brinda también la oportunidad de discutir y plantear posibles explicaciones a otras cuestiones fundamentales sobre la memoria, como el hecho de que sea constructiva (y no acumulativa), dinámica (que se encuentra en constante cambio) u orientada a la acción.

Un segundo rasgo que subyace a este texto, no menos importante que el anterior (y en realidad profundamente ligado a este), es la defensa de la memoria como una habilidad situada, lo que significa que los seres humanos somos agentes cognitivos actuando en un ambiente físico y social, y que nuestros procesos cognitivos se desarrollan, se desempeñan y se perfeccionan en comunicación y retroalimentación con ese ambiente (ver Barsalou, 2008; Fuchs, 2012; Rietveld, Denys & Van Westen, 2018).

---

22 Un tema de continuo debate es la pregunta de si existe un encapsulamiento de los procesos perceptivos de bajo nivel (como la percepción) que los protege de la interferencia de las funciones cognitivas de alto nivel (como memoria o lenguaje). Sin embargo, la investigación empírica ha fortalecido cada vez más el planteamiento de que el sistema cognitivo humano posee una arquitectura interactiva donde las funciones de alto nivel pueden influenciar la cognición de bajo nivel (véase Firestone & Scholl, 2015; Lupyan & Clark, 2015), lo que se ha conoce como permeabilidad cognitiva.

23 Como señalan los teóricos de la entrevista en historia oral Aron-Schnapper y Hanet (1993), las entrevistas "recogen de manera privilegiada cierto tipo de información en ciertos medios sociales y en ciertas situaciones [...] se emplean con el propósito de conservar y transmitir las historias [...] y, en general, todo lo que por su carácter no deja huella o deja pocas" (p. 63).

Explorar los factores contextuales que influyen el proceso de evocación autobiográfica, permite ver que los sistemas cognitivos humanos interactúan con el medio natural y social de forma constante; y que distintos elementos de esos medios afectan −y posiblemente forman parte constitutiva de− procesos cognitivos como la evocación.

Por último, es necesario mencionar que el presente artículo busca inscribirse en la discusión sobre estudios de memoria humana como una aportación teórica que reflexiona sobre un proceso complejo y multidimensional, como es la evocación. Si bien la ambición que lo motiva ha sido describir los procesos cognitivos que esta conlleva y mostrar la relación existente entre percepción y memoria durante su ejecución, el trabajo que se ha llevado a cabo pertenece a la esfera de la reflexión teórica y muchas de las afirmaciones que se expresan, a pesar de estar fundamentadas y referenciadas en trabajos de orden experimental, precisan de un sustento empírico mayor para ser contrastadas y verificadas. Sirva este trabajo, pues, de plataforma y punto de partida para futuras investigaciones interdisciplinarias que permitan ampliar el fundamento empírico de las ideas aquí ensayadas.

Referencias

Actis, N., Aldini, C., Gardella, L., Lewin, M. & Tokar, E. (2006). *Ese infierno. Conversaciones de cinco mujeres sobrevivientes de la ESMA*. Altamira.

Arfuch, L. (1995). *La entrevista, una invención dialógica.* Paidos.

Arfuch, L. (2005). Problemáticas de la identidad. En L. Arfuch (Comp.), *Identidades, sujetos y subjetividades* (pp. 21-44). Prometeo Libros.

Arfuch, L. (2007). *El espacio biográfico. Dilemas de la subjetividad contemporánea.* FCE.

Aron-Schnapper, D. & Hanet, D. (1993). De Herodoto a la grabadora: fuentes y archivos orales. En J. Aceves Lozano (Comp.), *Historia oral* (pp. 60-82). Inst. Dr. José María Mora.

Atkinson, A. & Adolphs, R. (2011). The neuropsychology of face perception: beyond simple dissociations and functional selectivity. *Philosophical Transactions of the Royal Society B, 366,* 1726–1738. https://doi.org/10.1098/rstb.2010.0349

Barela, L., Miguez, M. & García, L. (2009). *Algunos apuntes sobre Historia oral y cómo abordarla.* Dirección General Patrimonio e Instituto Histórico.

Baddeley, A. (1990). *Human memory: theory and practice.* Allyn & Bacon.

Baddeley, A., Eysenck, M. & Anderson, M. (2020). *Memory.* Routledge.

Barsalou, L. (2005). Continuity of the conceptual system across species. *Trends in Cognitive Sciences, 9*(7), 309-11. https://doi.org/10.1016/j.tics.2005.05.003

Barsalou, L. (2008). Grounded cognition. *Annual Review of Psychology, 59,* 617-45. https://doi.org/10.1146/annurev.psych.59.103006.093639

Boroditsky, L., Ham, W. & Ramscar, M. (2002). What is universal in event perception? Comparing English & Indonesian speakers. En W. Gray & C. Schunn (Eds.), *Proceedings of the twenty-fourth annual conference of the Cognitive Science Society* (pp. 136-141). https://doi.org/10.4324/9781315782379-61

Chomsky, N. (2006). *Language and mind.* Cambridge University Press.

Conway, M. (2001). Sensory-perceptual episodic memory and its context: autobiographical memory. *Philosophical Transactions of the Royal Society B, 356,* 1375-1384. https://doi.org/ 10.1098/rstb.2001.0940

Conway, M. (2003). Cognitive-affective mechanisms and processes in autobiographical memory: Commentary. *Memory, 11,* 217-224. https://doi.org/10.1080/741938205

Conway, M. (2005). Memory and the self. *Journal of Memory and Language, 53*(4), 594-628. https://doi.org/10.1016/j.jml.2005.08.005

Cueto, S. (2008). *Nacimos en su lucha, viven en la nuestra. Identidad, justicia y memoria en la agrupación HIJOS-La Plata.* Trabajo final de grado. Universidad Nacional de La Plata. Facultad de Humanidades y Ciencias de la Educación. http://www.memoria.fahce.unlp.edu.ar/tesis/te.426/te.426.pdf

De Brigard, F. (2014). Is memory for remembering? Recollection as a form of episodic hypothetical thinking. *Synthese, 191,* 155-185. https://doi.org/10.1111/phc3.12133

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58*(1), 17–22. https://doi.org/10.1037/h0046671

De Garay-Arellano, G. (1999). La entrevista de historia oral: ¿monólogo o conversación? *REDIE. Revista Electrónica de Investigación Educativa, 1*(1), 81-89. http://www.redalyc.org/pdf/155/15501107.pdf

De Garay-Arellano, G. (2001). Arquitectura de primera contra ciudadanía de segunda. Ponencia presentada en el *XIII International Congress of the Latin American Studies Association.* Washington DC, Septiembre 6-8.

Feldman, L., Mesquita, B. & Gendron, M. (2011). Context in emotion perception. *Current Directions in Psychological Science, 20*(5), 286-290. https://doi.org/10.1177/0963721411422522

Fisher, J. (2006). Does simulation theory really involve simulation? *Philosophical Psychology, 19*(4), 417-432. https://doi.org/10.1080/09515080600726377

Firestone, C. & Scholl, B. (2015). Cognition does not affect perception: Evaluating the evidence for "top-down" effects. *The Behavioral and Brain Sciences, 39,* e229. https://doi.org/10.1017/S0140525X15000965

Fuchs, T. (2012). The phenomenology of body memory. En S. Koch, T. Fuchs, M. Summa & C. Muller (Eds.), *Body Memory, Metaphor and Movement* (pp. 9-23). John Benjamins. https://doi.org/10.1075/aicr.84.03fuc

González-Monteagudo, J. (2010). La entrevista en Historia oral e Historias de vida: Teoría, método y subjetividad. En L. Benadiba (Ed.), *Historial oral: fundamentos metodológicos para reconstruir el pasado desde la diversidad* (pp. 21-38). Sur América Ediciones.

Gomila, A. (2012). Verbal minds. *Elsevier.* https://doi.org/10.1016/C2010-0-67019-2

Guerrero, C., Meilán, J., Carpi, A. & Palmero, F. (2008). Emoción y memoria: influencia del estado afectivo en el proceso de reconocimiento de palabras. En I. Etxebarría, A. Aritzeta, E. Barbera, M. Choliz, M. P. Jiménez, F. Martínez, P. Mateos & D. Páez (Eds.), *Emoción y motivación: contribuciones actuales* (pp. 109-118). Mitxelena

Guerrero Velásquez, C. A. (2020). Factores contextuales que influyen en la evocación autobiográfica en un contexto de entrevista [Tesis de maestría, Universidad Autónoma del estado de Morelos]. http://riaa.uaem.mx/handle/20.500.12055/1510

Grele, R. (1991). La Historia y sus lenguajes en la entrevista de historia oral: Quién contesta a las preguntas de quién y por qué. *Historia y Fuente Oral, 5,* 111-129. http://www.jstor.org/stable/27753314

Guttfreund, D. (1990). Effects of language usage on the emotional experience of Spanish-English and English-Spanish bilinguals. *Journal of Consulting and Clinical Psychology, 58,* 604-607. https://doi.org/10.1037/0022-006X.58.5.604

Hinojosa Luján, R. (2013). La historia oral y sus aportaciones a la investigación educativa. *Revista de Investigación Educativa de la REDIECH, 3*(5), 57-65. http://www.redalyc.org/articulo.oa?id=5216/521652343007

Javier, R., Barroso, F. & Muñoz, M. (1993). Autobiographical memory in bilinguals. *Journal of Psycholinguistic Research, 22*(3), 319-338. https://doi.org/10.1007/BF01068015

Keltner, D. & Ekman, P. (2000). Facial expression of emotion. En M. Lewis & J. Haviland-Jones (Eds.), *Handbook of Emotions* (pp. 236-249). Guilford Publications, Inc.

Lupyan, G. & Clark, A. (2015). Words and the world: Predictive coding and the languageperception-cognition interface. *Current Directions in Psychological Science, 24,* 279–284. https://doi.org/10.1177/0963721415570732

Marcos, L. (1976). Bilinguals in psychotherapy: Language as an emotional barrier. *American Journal of Psychotherapy, 30,* 552-560. https://doi.org/10.1176/appi.psychotherapy.1976.30.4.552

Marian, V. & Neisser, U. (2000). Language-dependent recall of autobiographical memories. *Journal of Experimental Psychology: General, 129*(3), 361-368. https://doi.org/10.1037//0096-3445.129.3.361

Michaelian, K. (2016). *Mental time travel: episodic memory and our knoledge of the personal past.* The MIT Press.

Michaelian, K. & Robins, S. (2018). Beyond the causal theory? Fifty years after Martin and Deutscher (13-32). En K. Michaelian, D. Debus & D. Perrin (Eds.), *New Directions in the Philosophy of Memory* (pp. 13-32). Routledge.

Michaelian, K. & Sant'Anna, A. (2019). Memory without content? Radical enactivism and (post) causal theories of memory. *Synthese.* https://doi.org/10.1007/s11229-019-02119-7

Mortensen, L., Berntsen, D. & Bohn, O. (2015). Retrieval of bilingual autobiographical memories: effects of cue language and cue imageability. *Memory, 23*(2), 138–156. https://doi.org/10.1080/09658211.2013.873809

Ortí, A. (1993): La apertura y el enfoque cualitativo o estructural: la entrevista abierta semidirectiva y la discusión de grupos. En M. García Ferrando, F. Alvira & J. Ibañez (Comps.), *El análisis de la realidad social: Métodos y técnicas de investigación* (pp. 171-203). Alianza.

Pezdeck, K. (2003). Event memory and autobiographical memory for the events of September 11, 2001. Applied Cognitive Psychology, *17,* 1033–1045. https://doi.org/10.1002/acp.984

Pinker, S. (1994). *The language instinct.* Penguin.

Pinto, C. (2013). *Los hijos de los exiliados vuelven a Chile: dilemas y desafíos para la integración memoria e identidad.* Tesis de posgrado. Universidad Nacional de La Plata. Facultad de Humanidades y Ciencias de la Educación. http://www.memoria.fahce.unlp.edu.ar/tesis/ te.883/te.883.pdf

Portelli, A. (2017). El uso de la entrevista en la Historia oral. *Anuario de la Escuela de Historia, 20,* 35-48. http://www.anuariodehistoria.unr.edu.ar/ojs/index.php/Anuario/article/view/205/224

Reese, E. & Fivush, R. (1993). Parental styles of talking about the past. *Developmental Psychology, 29*(3), 596-606. https://doi.org/10.1037/0012-1649.29.3.596

Rietveld, E., Denys, D. & Van Westen, M. (2018). Ecological-enactive cognition as engaging with a field of relevant affordances: the skilled intentionality framework (SIF). En A. Newen, L. De Bruin & S. Gallagher (Eds.), *The Oxford Handbook of 4E Cognition.* Oxford. https://doi.org/ 10.1093/oxfordhb/9780198735410.013.3

Robin, R. (1996). *Identidad, memoria y relato. La imposible narración de sí mismo.* Universidad de Buenos Aires.

Roediger, H. & Mcdermott, K. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(4), 803–814. https://doi.org/10.1037/0278-7393.21.4.803

Ross, M., Xun, W. Q. E. & Wilson, A. E. (2002). Language and the bicultural self. *Personality and Social Psychology Bulletin, 28,* 1040-1050. https://doi.org/10.1177/01461672022811003

Shanton, K. & Goldman, A. (2010). Simulation theory. *WIREs Cogni Sci, 1,* 527-538. https://doi.org/10.1002/wcs.33

Schacter, D. (1999). The seven sins of memory. Insights from psychology and cognitive neuroscience. *American Psychologist. 54*(3), 182-203. https://doi.org/10.1037/0003-066X.54.3.182

Schacter, D. (2002). *The seven sins of memory: how the mind forgets and remembers.* Houghton, Mifflin and Co.

Schacter, D. (2018). Implicit memory, constructive memory, and imagining the future: a career perspective. *Perspectives on Psychological Science, 14*(2), 256-272. https://doi.org/10.1177/1745691618803640

Schacter, D. & Addis, D. (2007). The cognitive neuroscience of constructive memory, remembering the past and imagining the future. *Philosophy Transactions of the Royal Society B, 362,* 773–786. https://doi.org/10.1098/rstb.2007.2087

Smith, E. & Kosslyn, S. (2008). *Procesos cognitivos: modelos y bases neurales* (M. Ramos Platón, Trad.). Pearson Educación.

Szpunar, K., Spreng, N & Shacter, D. (2014). A taxonomy of prospection: introducing an organizational framework for future oriented cognition. *Procedings of the National Academy of Science of the U S A, 111*(52), 18414–18421. https://doi.org/10.1073/pnas.141714 4111

Tsao, D. & Livingstone, M. (2008). Mechanisms of face perception. *Annual Reviews Neuroscience, 31,* 411–37. https://doi.org/10.1146/annurev.neuro.30.051606.094238

Tulving, A. & Craik, F. (2000). *The Oxford Handbook of Memory.* Oxford University Press.

Tulving, E. (2002). Episodic memory: from mind to brain. *Annual Review of Psychology, 53,* 1-25. https://doi.org/10.1146/annurev.psych.53.100901.135114

Whorf, B. (1956). *Language, thought & reality.* The MIT Press.

Estud.filos n.º 64. Julio-diciembre de 2021 | pp. 21-45 | Universidad de Antioquia | ISSN 0121-3628 | ISSN-e 2256-358X

45

https://doi.org/10.17533/udea.ef.n64a02

ARTÍCULOS

No maps for these territories: exploring philosophy of memory through photography*

Alun Kirby

York, United Kingdom
E-mail: streptococcus@gmail.com

Abstract: I begin by examining perception of photographs from two directions: what we think photographs are, and the aspects of mind involved when viewing photographs. Traditional photographs are shown to be mnemonic tools, and memory identified as a key part of the process by which photographs are fully perceived. Second, I describe the metamorphogram; a non-traditional photograph which fits specific, author-defined criteria for being memory. The metamorphogram is shown to be analogous to a composite of all an individual's episodic memories. Finally, using the metamorphogram in artistic works suggests a bi-directional relationship between individual autobiographical memory and shared cultural memory. A model of this relationship fails to align with existing definitions of cultural memory, and may represent a new form: sociobiographical memory. I propose that the experiences documented here make the case for promoting a mutually beneficial relationship between philosophy and other creative disciplines, including photography.

Keywords: photography, memory, cultural memory, perception, art

No hay mapas para estos territorios: una exploración de la filosofía de la memoria a través de la fotografía

Resumen: Al inicio se examina la percepción de las fotografías desde dos direcciones: qué pensamos que son las fotografías y los aspectos de la mente involucrados al ver fotografías. Se muestra que las fotografías tradicionales son herramientas mnemónicas: por ello, la memoria es una parte clave del proceso mediante el cual las fotografías se perciben por completo. En segundo lugar, describo el metamorfograma; una fotografía no tradicional que se ajusta a criterios específicos definidos por el autor para ser memoria. Se muestra que el metamorfograma es análogo a una combinación de todos los recuerdos episódicos de un individuo. Finalmente, el uso del metamorfograma en obras artísticas sugiere una relación bidireccional entre la memoria autobiográfica individual y la memoria cultural compartida. Un modelo de esta relación no se alinea con las definiciones existentes de memoria cultural y puede representar una nueva forma: la memoria sociobiográfica. Propongo que las experiencias aquí documentadas justifican la promoción de una relación mutuamente beneficiosa entre la filosofía y otras disciplinas creativas, incluida la fotografía.

Palabras clave: fotografía, memoria, memoria cultural, percepción, arte

Alun Kirby

was an immunologist for many years, researching immune responses to mucosal infections in London, Sweden, and York, and making photographs all the while. Since leaving science he has focused on artistic work, using philosophy of memory and the lived experiences of people living with dementia as sources of inspiration.

ORCID: 0000-0001-5663-2961

Introduction

How do we perceive photographs? Can a photograph ever be a memory? What can we learn about memory, and our relationship with memory, from photographs? I have been exploring philosophical ideas of memory and its relationship with the self through photography, and this paper sets out the lessons learnt during that process.

In the first section I examine our relationship with photographs. The culturally accepted analogy equating photographs with memories is contrasted with their experimentally observed functions. The actual roles photographic images play in relation to memory defines them not as 'memories' but as mnemonic tools. I then examine the role of memory in our perception of photographs by revisiting the work of Roland Barthes. By placing his documented experience in the context of contemporary theories, the potential contribution of memory to affective perception of photographs is revealed.

Typically, photographs are second-generation copies of visual scenes external to the camera, captured in a fraction of a second. However, other photographs can be created that do not require a camera and do their work over long time periods. These photographs do not record phenomena external to themselves, instead recording what they directly 'perceive'. The second section explains how I developed a photographic process broadly analogous to physical human memory: the metamorphogram. The type of memory represented by metamorphograms is examined against existing definitions of procedural, semantic and episodic memories, showing most similarity to a composite of all episodic memories recorded during an individual's lifetime.

In the final section, this photographic analogy is used to explore relationships between individual and cultural/collective/social memory. Inspired by a Japanese folktale, several works were created which implied a bi-directional relationship between memory and self: that individuals can alter culturally shared memories, and in return cultural memories can affect individuals at the level of their autobiographical self. This relationship is explored in more detail, revealing a distinct form of non-institutionalised, non-behavioural, socially shared memory. Not fitting any extant definition, it is proposed that this be entitled 'sociobiographic memory'.

1. Our perception of photographs

'We capture your memories forever'

> 'We are that strange species that constructs
> artefacts intended to counter the natural flow
> of forgetting'
> (William Gibson, 2012)[1]

---

1   This quote is from William Gibson's collection of non-fiction writings, 'Distrust That Particular Flavour' (2012), at the start of the article *Dead man sings* (pp. 51–53), in which Gibson considers our rapidly changing relationship with the past.

The question 'how do we perceive photographs?' can be interpreted in two ways, depending on our definition of 'perceive'. First, it may mean 'what do we recognize or understand photographs to be?' As a society as well as individuals, we blithely entrust the recording of important memories to cameras; or, more precisely, to photographs. We even call the most perfect, albeit apparently non-existent, version of memory – in which every detail is recalled with absolute clarity– "photographic memory". A generalized acceptance exists within society that a photograph is somehow equivalent to a memory. This apparent equivalence has been promoted and exploited by camera manufacturers Eastman Kodak since the 1960s and 1970s onwards. One print advertisement copy includes the following;

> Then and Now Good Memories Deserve Good Processing
>
> What are memories made of? [...] You live them all once. You live them again and again in pictures [...]. And make your memories last.

As economist Theodore Levitt noted, "Kodak promises with unremitting emphasis the satisfactions of enduring remembrance, of memories clearly preserved. [...] The product is thus remembrance, not film or pictures."[2] The messages are clear: memories are 'captured' in photographs; photographs extend the functional lifespan of memories. The arts writer John Berger concurred, suggesting that photographs may be a direct replacement for memory:

> What served in the place of the photograph, before the camera's invention? The expected answer is the engraving, the drawing, the painting. The more revealing answer might be: memory (Berger, 1980, p. 54).

In this he iterated Susan Sontag, who stated that photographs are "not so much an instrument of memory as an invention of it or a replacement" (Sontag, 1979, p. 178). And so, we entrust our fallible memories to cameras which "cannot lie".

However, Kodak, Berger and Sontag are wrong; in fact, Sontag's quote works better reversed, to infer that photographs are not a replacement of memory, but an instrument of it. Photographs are not analogous to memories, but to tools. Participants in studies of how we use photographs directly refer to them as memories: "It's important for me that they'll have [the picture collection] when they are grown up, so they will be able to leave home with a big box of memories" (Whittaker, Bergman & Clough, 2010, p. 34). However, when participants' interactions with photographs are further interrogated, it is clear that they understand the functional purposes, whether memorial, communicative or aesthetic (Fawns, 2020, pp. 3, 6, 8 & 9), for which they use photographs. Fawns describes a multitude of ways in which photographs are used to

---

2   From the article '*Marketing intangible products and product intangibles*', Harvard Business Review, May–June 1981, pp. 94-102

support (autobiographical) memory construction, recall, association and distribution. Participants in Fawns' studies understand that they use photographs as tools to aid recall, noting that selecting images to keep is "important potentially in the future and you want to make the right choices" (Fawns, 2020, p. 6). They show understanding of how photographs trigger memory: "The photograph becomes the focal point of the memory system that everything then extends out from" (Fawns, 2020, p. 3). Fawns sees photographs as part of a process, including taking, sorting, viewing and distributing images, which we use to support memory. The photographs themselves are merely one of the tools in this process, interacting with our memory, functioning as symbols to be interpreted (Rowlands, 2016, pp. 160-162).

Our reliance on photographs as mnemonic tools is a double-edged sword. Photographs have been shown to aid reactivation of memories, but may also induce bias in our memory (see St Jacques & Schacter, 2013, pp. 537-343; St Jacques, Montgomery & Schacter, 2015, pp. 876-887 for examples), or, albeit in somewhat contrived circumstances, create completely false memories (Garry & Gerrie, 2005, pp. 312-315). Photographs clearly function in multiple and complex ways with our memory, often in the context of other contemporary non-visual (e.g. textual, verbal, social) information (reviewed in Garry & Gerrie, 2005). The strength of their potential influence led to the suggestion that photographs supporting memories can validate a sense of self (Heersmink & McCarroll, 2019, pp. 98-101).

Decades of research have provided objective clarity in terms of the function of photographs. Despite this, the power and simplicity of the label 'memory' for driving engagement with images (as evidenced by Instagram Memories, Google Photo Memories, Facebook Memories... and the continuing use of 'memories' as a hook word by almost every major camera manufacturer) suggests photographs will remain commonly perceived as memories. After all, who would buy a camera that lets you capture "visual tools to aid your recall"?

Punctum power

We have seen that the generalized perception of photographs as memories is at odds with their actual function as mnemonic tools. Given this role in triggering recall, I now return to the question 'how do we perceive photographs?'.

The primary definition of 'perceive' by the *Oxford English Dictionary* is 'to take in or apprehend with the mind or senses' (further: 'to apprehend with the mind; to become aware or conscious of; to realize; to discern, observe').[3] In this section, I restate the

---

3   The Oxford English Dictionary (online) was the source for definitions in this work. For the definition of 'perceive' see: https://www.oed.com/view/Entry/140537

question as 'with what aspects of our mind or senses do we interpret photographs?' To answer I focus on Roland Barthes' observations in *Camera Lucida* (Barthes, 1980). Barthes became fascinated by photography and photographs, examining the viewer's relationship with images. Barthes' testimony will be re-evaluated in the light of more contemporary philosophical theories.

In *Camera Lucida* Barthes describes finding a box of photographs belonging to his beloved mother, who had recently died. In some he recognizes his mother and other family members, with little affective consequence despite his obvious grief. However, in the "Winter Garden photograph", which shows his mother as a child, he finds something immediately and profoundly affective. He attests that this image has captured the essence of his mother, allowing him to 'rediscover' her. How has Barthes perceived this photograph, and not others of his mother, such that it affects him so deeply? The image is explicitly not an autobiographical memory: it shows his mother as a young girl, long before he was born. Indeed, Barthes knows that "what [he sees] is not a memory" (Barthes, 1980, p. 82) and that "The Photograph [is] never, in essence, a memory" (Barthes, 1980, p. 91). For him it is "reality in a past state"; confirmation that what he sees once existed and has "evidential force" (Barthes, 1980, p. 89).

Barthes found affective power in some photographs, and knows this power relates directly to the viewer. He says of the Winter Garden photograph, "[F]or you, it would be nothing but an indifferent picture" (Barthes, 1980, p. 73). What strikes Barthes about photographs with affective content is something he describes as an image's "intensity"; the part of the image which "jumps out and pierces" the viewer, an aspect of visual content which he calls a "punctum" and which "is poignant to [him]" (Barthes, 1980, pp. 26-27). His perception of the punctum is crucial for the effect, and Barthes understands that its presence or absence is viewer-specific.

Perceiving the punctum

Above, I defined the question under examination as "with what aspects of our mind or senses do we interpret photographs?". So far, we have Barthes viewing photographs, but only finding affective content in a fraction of them. This suggests that while perception of photographs includes input of visual information, other aspects of perception facilitate recognition of a viewer-specific punctum able to transform indifferent pictures into objects with the affective capacity.

What is this punctum? Baudrillard, another philosopher captivated by photography, said that in most photography "what Barthes calls the 'punctum', that absent point, that nothingness at the heart of the image which gives it its power, no longer exists". Baudrillard appears to equate the punctum with a "void": a complete absence of symbolism. He suggests that taking a photograph of a living human which

contains a punctum may be "almost impossible" (Baudrillard, 1998, p. 93) as there can be no "absence".

Barthes seems to say that the punctum is more obvious when Death is present in the image. In the Winter Garden photograph and other historical images he notes that the subject is at once alive / going to die / has died / dead, suggesting some necessary relation to time or, as Barthes melancholically phrases it, "death in the future" (Barthes, 1980, p. 94). However, by no means all the puncta he uses to illustrate this idea are people. The punctum may be an object, a stance, a house, a sheet, or the entirety of the image. A punctum has the "power of expansion", which Barthes says is often metonymic. He sees one thing, and his mind fills with much, much more. From a picture of a blind gypsy he "recognize[s], with [his] whole body" his long-past travels in Eastern Europe (Barthes, 1980, p. 45). In a photograph of an unknown woman he realises the punctum is due to his associating her necklace with that of someone he once knew (Barthes, 1980, p. 53), and from there a whole trail of memories related to this object unfolds within him. This same associationist phenomenon of "expansion" is also noted by participants in recent studies (Fawns, 2020: see the first section for an example).

The power of the punctum thus emanates from the viewer. Barthes and I would not share puncta, we would each find our own. I propose that each punctum's power is due to a second aspect of our perception of photographs; that we perceive photographs in the context of our memory. The consequence of this perception is some level of emotional affect.

Not all images contain a punctum; they show nothing which "pierces" our memory and induces affect. Many are "inert under [Barthes'] gaze" (Barthes, 1980, p. 27), others elicit semantic memories such as recognition of a particular form of clothing. Photographs may therefore act with a scale of affect, ranging from nothing at all, through factual recognition, and all levels of "punctum" affect up to a punch-in-the-gut emotive force. The viewer-specific power of a photograph arises from the association of something within the image with an aspect of the viewer's memory.

In the Winter Garden photograph, Barthes' perception of the punctum which provides the affect is clearly expressed. Barthes receives the visual stimulus of his mother as a child, sees her particular pose, the line of her face, and is filled with emotion. The image itself has no intrinsic power; I doubt you or I would be affected should we see it. Barthes' memory has been triggered to engage with this photograph. Therefore, the viewer's memory, in concert with visual stimuli, determines whether any personally relevant or affective content is perceived.

From Barthes' account we see the scale of affect puncta may trigger. Only the Winter Garden photograph arouses the notable affective state he discusses, despite looking at other photographs of his mother from the same box, including one "in which she is hugging me, a child, against her, [and] I can awaken in myself the rumpled softness of her crepe de Chine and the perfume of her rice powder" (Barthes, 1980, p. 65). Degree of affect can thus distinguish one image from another, even when the representational

Estud.filos n.º 64. Julio-diciembre de 2021 | pp. 47-71 | Universidad de Antioquia | ISSN 0121-3628 | ISSN-e 2256-358X

53

https://doi.org/10.17533/udea.ef.n64a03

content —in this case, Barthes' mother— is very similar. The Winter Garden photograph may be set apart as particularly valued for its ability to conjure such strongly emotional memories. Or, we may say, for the ability of Barthes' memory to facilitate perception of it in such an intensely affective manner.

Barthes' response to the Winter Garden photograph is so strong we may consider it as an evocative object, one which "intentionally or unintentionally aids us in remembering our personal past" (Heersmink, 2020, p. 7). The intensity of Barthes feeling suggests a "love-at-first-sight" moment, a viscerally affective flood of connected memories triggered by the image. Barthes could be describing the moments during which the Winter Garden photograph is transformed, for him, from a mere image into what we might now label a particularly powerful evocative object.

Overall, Barthes' testimony, together with more recent empirical observations, suggest that perception of photographs is a complex process in which visual stimuli interact with memory to facilitate perception of anything of relevance to the viewer within the image, be that purely semantic information or associations with affective autobiographical content. All this is done in the context of the viewer's present cultural, social and personal situation, rather than the (distant or recent) past depicted in the image. It is within this network of interactions that our mind fully perceives a photograph.

2. The metamorphogram

Photographic memory

I have shown that photographs do not function as memory. However, so far we have only looked at "traditional" photographs. By this I mean images of scenes external to the camera or film, which exist only for the fraction of a second it takes for the shutter to open and close —usually less time than our brains are capable of perceiving. My artistic practice uses non-traditional photographic techniques, which I have previously used to create visual analogies for philosophical ideas. For example, a photographic sculpture constructed along associationist principles,[4] or a series of images using the same source material to depict the reconstructive nature of memory.[5] These explorations led to the question 'can a photograph ever be a memory?' That is, is it possible to create photographs which are directly analogous to a memory?

I will first define "a memory" with relatively broad yet consciously restrictive terms chosen to replicate, or be representative of, central characteristics of human

---

4   For details on my sculptural piece please see: 'Associationism' http://alunkirby.com/portfolio/reconstructed-memory-associationism/

5   '(Re (Re)) Reconstructed Memory was exhibited in 2020 as part of this exhibition: 'False Memory' https://www.ragm.co.uk/falsememory

episodic, autobiographical memory. They are designed to allow easy comparisons between traditional photography and the approach described below, and to allow flexibility within a practical artistic context. Such a memory should: (1) be created as a result of direct perceptual experience, (2) be recorded directly and independently by the object having the experience, and (3) be recallable. The criterion of "direct perceptual experience" is intended to push away from semantic into episodic and autobiographical areas of memory. In addition, a photographic analogy of a memory should (4) function on the same timescale as human episodic memory, to ensure the capture of an experience, rather than a transient external scene. Traditional photographs do not fit these criteria; we may exclude them at criterion (1) by noting that such images are indirectly experienced. The exposure to the scene visualized happened to the film or digital sensor; traditional photographs are secondary interpretations.

To fulfil these criteria and approach this analogy photographically I used a process called cyanotype (blueprint), which uses light-sensitive chemicals and has no requirement for a camera.[6] Photograph means "light drawing", and many of the earliest photographs were cameraless. To create a cyanotype, a photosensitive solution is usually painted onto paper. When exposed to ultra-violet light (sunshine, for example), a permanent blue colour is formed. In contrast to the fractions of a second over which traditional photographs work, cyanotypes can be left to expose for months if needed. This allows the cyanotype to fulfil criterion (4), working on a human-relatable timescale. To fulfil criteria 1-3, I chose to create an approximation of human physical memory. Our capacity to record and recall is held in our brain, contained in our body, with the ability to perceive external stimuli through our sensorium. Through these physical components our experiences are processed, recorded, and made available for recall.

The 'brain' of the photograph is the photo-sensitive solution applied to the paper, providing the record-and-recall function. The 'body' is the paper itself, which can be given almost any shape we wish through folding. The surface of the paper capable of being exposed to light is analogous to the sensorium, which 'responds' to external stimuli by altering the chemicals.

Finally, a critical aspect of this attempt to create a photographic model of an episodic-type memory is that it should result from direct experience. The definition of experience in this case is "an event by which one is affected"; a state or condition to which one is subject. To create an authentic episodic memory I experience an event directly, and translate my perceptions of the event into a memory of that event. The model of creating a cyanotype-memory, below, is based on this human-level process.

---

6   Cyanotype is an iron-based, simple photographic process invented in 1842. Cyanotype was used by Anna Atkins to create the first ever book containing photographs, which can be viewed at the British Library; https://www.bl.uk/catalogues/photographyinbooks/record.asp?RecordID=3048

Making the metamorphogram

Here I describe the process used as an analogy of human physical form capable of recording its own episodic memories. Paper coated with light-sensitive cyanotype chemicals provides the capacity to record and recall external visual stimuli –the brain of the model (Figure 1a).

(a)   **(b)**   **(c)**
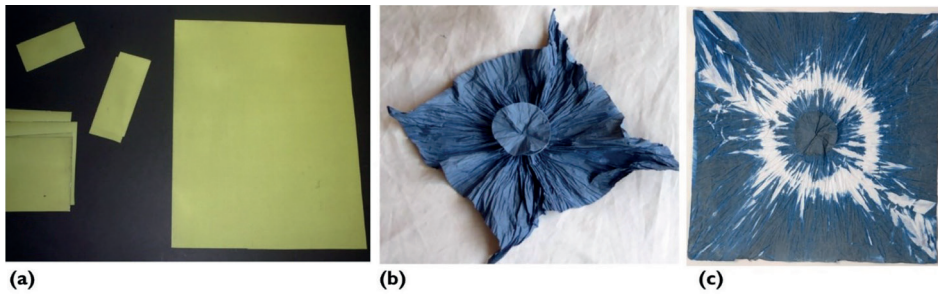
Figures 1a, b & c
Production of a metamorphogram (2019)

(a) Paper coated with photosensitive chemicals prior to use. (b) An abstract 3D form is folded and allowed to expose. (c) The 'killed' final form, showing the resultant image. In white areas no light reached the paper. Blue areas were exposed to sunlight. Greenish-blue areas, such as the central circle, received most light.

The paper is then altered, usually using origami techniques, to give a form specific to the individual model (e.g. Figure 1b). Forms were either nominally two-dimensional, essentially various layers of paper, or three-dimensional, where the form is more-or-less limitless. Now we have a 'body' (the form), a 'brain' (the light-sensitive chemicals), and the primarily outward-facing, chemical-coated surface(s) of the form comprise the 'sensorium' (areas capable of perceiving and responding to external stimuli). It may be helpful to consider the form at this stage as analogous to a new-born creature, with no autobiographical memory, ready to have experiences and create its own episodic memories.

The process is not intended to create an equivalent to human memory, but a small-scale analogy of it, with a sensorium and capacity to record and recall information limited to certain visual stimuli. To complete the process and determine whether a photograph can really fulfil the criteria for being a memory it needs to have experience(s) –events by which it is affected– and record and recall them.

Experience was not dictated. That is, apart from early "proof-of-principle" studies, in which the events were somewhat controlled, the locations where forms were placed and what happened to them during their "life" was left to happenstance, and forms were treated arbitrarily. Once created they are placed into the environment (usually inside my home) and may be moved or left still, played with or ignored, forgotten … neither rhyme nor reason dictated the 'events-which-affect-them', their experience.

Accumulation of experiences was allowed to continue until some arbitrary external force ended it. As creator, I acted as 'arbitrary external force'. Some forms lived a few hours, some days, weeks or months. Some had their own adventures (were lost) until the arbitrary external force caught up with them. Through these methods it was intended that every individual form had unique experiences.

Once the experiential time is over, the form is 'killed', which is merely the undoing of form, returning it to a relatively flat piece of paper. At this stage any "memory" is statically recorded in chemical changes which have occurred within the surface of the paper. To complete the process, the "memory" is recalled by washing the paper in water to remove any unused chemicals and leave an extremely light- and colour-fast final image (Figure 1c). I have named the photograph created by this process —transforming paper into a distinct form which then returns to a flat image— a metamorphogram.

Experience-dependent photographic memory

We now have a photographic process adapted to create a loose analogy of autobiographical memory as instantiated in the human body. Forms are allowed to have 'subjective' experiences, events-which-affect-them, which creates the final image.

The following simple demonstration examines the crucial formative role of direct subjective experience in creating the final image. Two identical pieces of tissue paper were coated with the same batch of chemicals at the same time, and folded into exactly the same form; a flat origami called a "tato", which has all its folds on one surface (Figure 2a). These were placed in sunlight adjacent to one another for the same length of time; in this case several days. The only difference between the two forms is that one was placed folded-side up, while the other is folded-side down. Their creation and "life experience" is identical in every way, with "killing" and recall also identical in each case.
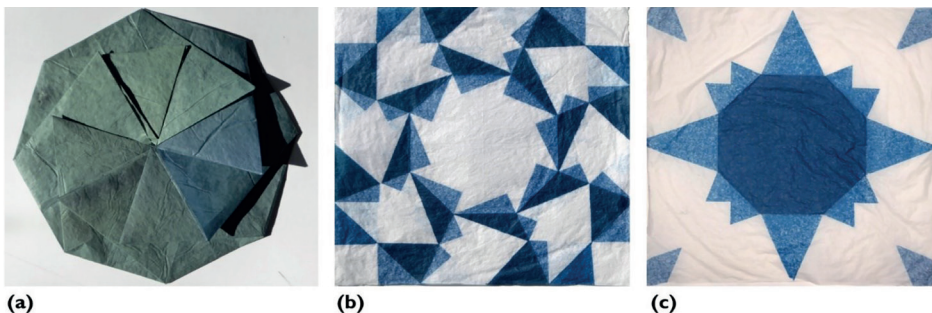
Figures 2a, b & c

The influence of experience on the visual image recorded by a metamorphogram (2017)

Coated tissue paper was folded into the tato form (a) and left to gain experience either folds-up (b) or folds down (c). All other variables were identical. (b) and (c) are the final 'killed' images from each form.

The images produced are dramatically different (Figures 2b and 2c). We can say one is the experience of 'being-a-tato-face-up' and the other the experience of 'being-a-tato-face-down'. Their single difference in perceptual experience directly and strikingly affected the image recorded of their otherwise identical 'life-as-a-tato'. We may imagine the analogous situation of two people standing back-to-back; they would have different perceptual experiences, and therefore different memories, of a particular event.

We are now able to ask whether these finished metamorphogram images fit the previously given definition of a memory. Are they (1) created as a result of direct perceptual experience? Yes, and the above experiment shows that the subjective nature of the experience affects the final image "memory". Was the image (2) recorded directly and independently by the object having the experience? Yes, as no external force, device or process was required for recording the image (traditional photography normally requires a camera and human operator or other physical mechanism to choose and trigger exposure and recording). Was the recorded information (3) recallable? Yes, we may consider this as analogous to me asking somebody to recall an event. I trigger and perceive the recalled memory, but the recall is independent from me. For a metamorphogram, I 'ask' it, for example, what it was like to 'be-a-tato-face-up'. My external stimulus 'initiates' the recall (and in this case physically aids recall by washing the paper), but the memory recalled is entirely specific to, and dependent on, what the object had recorded during the experience. Finally, the image was (4) created over a human-relatable time period, and so is much more reflective of human memory than the single instant recorded in a traditional photograph. This allows the conclusion that a metamorphogram image is a photograph which is directly analogous to a memory, based on these criteria.

Metamorphogram memory

I will now give further examples showing how the images created relate directly to the perceptual information recorded during an individual object's experience. We must note that metamorphograms are highly limited in both their perceptual and recording capabilities. They can only record using the chemicals on their surface and, to some extent, perhaps with their physical 'body'; the paper from which they are formed.

In a project working with people living with dementia, 3D origami shapes made from chemical-coated paper were fitted with recording devices.[7] Participants could tell their thoughts to the boxes, which would 'remember' what they said, while the box itself

---

created its own memory of its time (weeks to months) living with the person. Some of the images are shown in Figure 3.

With practical understanding, these metamorphograms can be 'translated' in the context of the known experience. For example, in Figure 3a:

- the visible folds and the pattern of coloured areas show this was a cube,

- the overall colour has become a gold-green, which indicates that this box was 'alive' in a bright place for several weeks at least,

- the far right, darkest in colour, was facing the light,

- the triangles below the gold squares are much paler than those above the squares due to experiencing much less light, and so formed the bottom of the box, and

- most interestingly, the random marks and smears in the coloured areas show us that this box lived in the participants kitchen, near the sink. The marks are splashes from the washing-up water which altered the chemicals in those places.

(a) **(b)** **(c)**
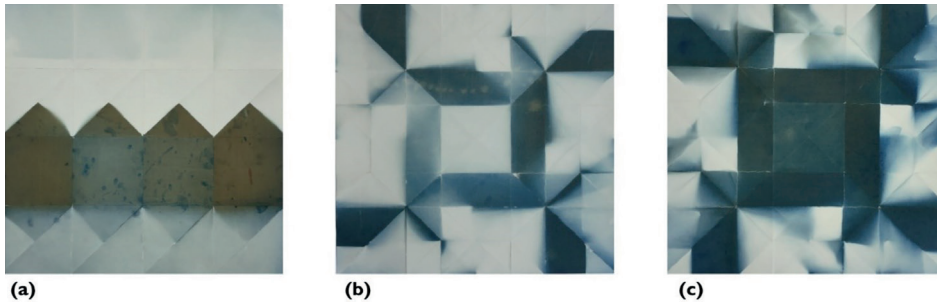
Figures 3a, b & c
Metamorphogram images from the Memory Box project (2018)

Box (a) has recorded both visual and physical stimuli, as well as multiple interactions
with its external environment. Boxes (b) and (c) show again the influence of experience
on the final image / memory produced.

Together these visual characteristics constitute the memory of this metamorphogram's 'life' as a box-form in this person's house. Unlike a traditional photograph it has not simply recorded a transient moment external to itself. It has, within the limits of its capacity to do so, recorded its direct perceptual experiences —formation, placement, duration, light exposure, orientation, and even the consequences of interactions with others.

In Figures 3b and 3c we see the visual memories of two boxes with the same initial form, but which spent their lives with different people. The box on the left (Figure 3b) was in a fairly bright place, with light coming mostly from the right, and shows a splash where some tea was spilt. The square in the centre was the bottom of the box, and so is white.

The box on the right (Figure 3c) shows more colour, and the colour is relatively even. It has had more light, and from all sides. Interestingly, the square in the centre is blue, which means the bottom of the box has been exposed to light. It happens that the participant moved house during the project, so the box experienced movement including being turned upside down for a while. The box recorded that experience, and our process has recalled the memory of it in visual form: we can visualise the biography of the form. This is a self-created biography, distinct from the biographies ascribed to personally or culturally important objects predominantly through human agency, as described by, for example, Hoskins (2005).

We seem to have a photograph which fulfils our criteria for memory. But what sort of memory? The analysis above also identified experience recorded in the physical media of the object itself, i.e. folds in the paper. Let us deal with this first.

We may argue that folds in the metamorphogram constitute procedural memory. The folds made to create the form remain, and often show which direction the paper was folded. A competent origamist could refold the form from these creases, but the paper could not refold itself, arguing against procedural memory. We may instead argue that folds are another visual aspect of our analogy to memory, but we have defined the paper as the form's 'body' rather than part of the light-sensitive sensorium. I conclude that these folds represent 'birth marks' or scars. They act as reminders of an experience, rather than a memory of that experience.

We are left with our photographic image. What kind of memory is it analogous to? As the image does not show non-autobiographical, non-experiential factual information, I exclude it being analogous to semantic memory. This leaves episodic memory. Some aspects of the metamorphogram relate to specific episodic experiences, such as the tea splash in Figure 3b. The washing-up splashes in Figure 3a are likely due to repeated incidents (note that the experience of being splashed is recorded by the photographic 'sensorium'; these marks are not mere 'stains' masquerading as a memory), and so result from multiple individual episodic experiences.

The colour of each image also arises due to the overlaid memory from multiple episodic experiences of 'being-in-a-particular-place' and recording the specific amount of light (dependent on direction, intensity etc.) perceived by the sensorium each time. A little colour from one dull day, overlaid with more colour from the next, brighter day, colour on different sides when moved for a day, and so on. Together these experiences result in an image which is a composite memory built up from many individual 'episodes'.

Metamorphograms are not, then, engrams or memory traces. They do not show the multiple aspects of a single experience, unless we consider their entire 'life-as-a-form' as a single experience. They are instead an amalgamation of all the recordable perceptual experience perceived by the form during its entire 'lifespan'. The analogy here would be to every episodic memory retained by a person, overlaid one on top of the next.

3. The individual and cultural memory

In this section I set out how artistic explorations led from initial examinations of individual memory to finding connections between individual and collective memory. I show that bi-directional agency exists between individuals and certain social groups to which they belong, such that (1) individuals can directly influence the nature of a collective memory, and (2) that collective memory can influence individuals at the level of their autobiographical 'self'.

I have shown that, while traditional photographs are tools rather than memories, it is possible to create a photographic image analogous to memory. These metamorphogram images appear to comprise a 'lifetime' of episodic memories in a single expression. Applying artistic licence, I now take a metamorphogram's memories at the time of its 'killing' to be somewhat comparable to an individual 'self', if it were possible to visualise such a thing. Each metamorphogram is unique due to its specific and subjective recorded experience, despite sharing features such as creative process, form and sensorium with other metamorphograms.

My interest in origami led me to explore what I considered to be 'cultural memory', long before I had read any philosophical or sociological definitions of what cultural memory might be. Naively, I thought the practice of origami may itself comprise a cultural memory. These early inquiries coalesced into one two-pronged question: How does cultural memory affect the individual, and how does the individual affect cultural memory?

At this stage, I had no definition of cultural memory in mind. However, in deciding that I could consider each metamorphogram as representing a unique individual, I was able to explore the relationship between cultural and individual memory through art.

There is a Japanese folk tale surrounding the origami crane, known as the orizuru (ori – paper; tsuru – crane). The red-crowned crane, as a bird, has various symbolic meanings in Japan, including longevity and authority. The origami version long predates its appearance in the first printed origami book, *Hiden Senbazuru Orikata* of 1797 —origami has existed in Japan since around 600AD.[8] Pre-second world war, the act of folding 1000 orizuru— called senbazuru —was thought to bring the folder a boon, such as long life, good health, or a wish. After the atomic bomb was dropped on Hiroshima, a girl called Sadako Sasaki developed leukaemia. She decided to fold senbazuru to wish away her illness, but died while still folding them (Sasaki & DiCicco, 2020). As a result of the perception of her experience by others, the cultural meaning of senbazuru has been altered. It has become

---

8   Much information can be found online regarding the history of origami. For example: https://en.wikipedia.org/wiki/History_of_origami includes mention of *Hiden Senbazuru Orikata*, which can also be viewed in full here: http://www.origamiheaven.com/senbazuruorikata.htm

a symbol of hope and peace; hundreds of thousands of cranes are delivered to shrines at Hiroshima and Nagasaki every year. Senbazuru projects for peace are now abundant across the world.

This story led me to construct two hypotheses. First, that an individual could directly affect cultural memory; in this case, expanding or altering the meaning of a long-standing cultural symbol. Second, that some cultural memories could affect individuals at the level of their autobiographical 'self'; people temporally and spatially distant from the original event are affected enough to fold one thousand cranes, use the imagery and symbolism, and encourage others to do so – enough to become a lesser or greater part of 'who they are'.

To test these hypotheses through art I decided to fold my own senbazuru, where each of the 1000 cranes would become a metamorphogram.[9] That is, cranes were folded from chemical-coated paper, allowed to 'live' according to the principles set out in section 2, above, and then 'killed' to reveal the the visual memory of their 'time-as-an-orizuru'. Although each of the 1000 metamorphograms was formed in an identical process, each was also unique as a final image (Figure 4a), representing the recorded experiences of a unique individual, and many of those experiences were shared across the 1000 individuals. I therefore had a group of individuals with shared experiences analogous to some sort of 'cultural' memory, and began to explore the relationships between them.

In *Cliques and Networks* (Figure 4b) 81 crane metamorphograms were arranged together. The individual metamorphograms cannot communicate with each other, so the relationships formed are based entirely on the memories they recorded while 'alive' in crane form. The patterns which appear infer how our personal experiences lead us to form a range of connections with others, and together those interactions form something new. Small groups of individuals together create circular 'cliques' where aspects of their individual memories meet. Other parts of the individual images form an extended diagonal grid, or 'network', connecting individuals well beyond our immediate social circle or clique.

---

9   The project to fold 1000 orizuru and express them as metamorphograms was inspired by the folk tale boon of a wish, and so is entitled '1000th of a Wish'. More details can be seen here: 1000th of a Wish Project http://alunkirby.com/portfolio/one-thousandth-wish-senbazuru/
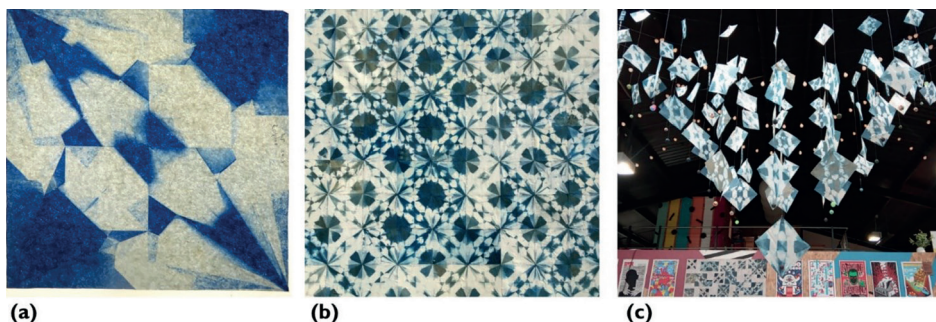
Figures 4a, b & c
*Exploring interactions between individual memory and collective
or cultural memory (2019-20)*

(a) An individual orizuru metamorphogram. (b) *Cliques & Networks* collage. (c) *Six Degrees of
Separation* installation of over 100 individual metamorphogram cranes.

The artwork visualizes through these linked individuals the concepts that each individual has its own memories; that some memories (i.e., experiences) are shared with a close social group, and other memories or experiences are shared on a much wider social level.

The second work shown is a hung installation entitled *Six Degrees of Separation* (Figure 4c), which has individual metamorphograms hung in five grouped layers, with the sixth layer to be provided by the viewer. The title arises from the popular, yet inaccurate, notion that any individual is connected to any other via no more than six social interactions. The five layers of individuals are 'connected' via their shared experience, which has resulted in them appearing very similar to one another. The work asks of the viewer, are you able to connect as the sixth layer, or do you lack a shared experience which would allow that connection? Individually recorded shared experience – a form of cultural or social memory – allows us to come together as a social group. The experiences gained within the social group may then feed back and become part of each individual's memory.

These works, created naively, without reference to literature on cultural memory, seemed to support the idea that a memory of an event experienced by a cultural or social group could link that group together, even though the event was experienced at an individual level. Further, that the coming together of individual memories created something new, new connections, or social structures. Further, that it was each individual's memory that combined to create these new things. The model represented by the artworks is certainly weak, almost barely there. However, it opened a pathway down which the idea developed.

Considering these works strengthened my conviction that the original linked hypotheses (1. that an individual can directly affect cultural memory, and 2. that

cultural memory can affect individuals at the level of their autobiographical 'self') were worthy of a more rigorous examination. To do this, two significant issues needed to be addressed. First, the art-based models only indirectly address the hypotheses. At this stage their validity is based on extrapolation and artistic inference. A stronger, real-world example is required. Second, no definition of 'cultural memory' has been given here. Since various types of cultural memory exist in philosophical and sociological literature (see, e.g., Adams, 2019; Manier & Hirst, 2008; Wang, 2021), any model needs to be tested against them.

City or United?

While the senbazuru example offered a glimpse of something, a more concrete example is needed to fully examine whether there exists a bidirectional flow of influence between individual and what, lacking a more specific term at this point, I will continue to call cultural memory until I compare it with extant models.

The hypothesis indicates that I should seek a social or cultural context to which individuals contribute, and by which individuals are influenced at the level of their autobiographical memory or 'self'. I choose to use as my model supporters of English Premier League Football clubs, although many other organisations, clubs or social groups would substitute.

Major football clubs are large businesses, their identity and product branded and sold around the world. However, football clubs require supporters, and the supporters' experience of football is far from institutionalized, and extremely social. Of importance for later comparisons with other forms of cultural memory, relationships between individual fans do not require familial, employment or other common social links. While support for a specific club may 'run in the family', with successive generations supporting the same team, intrafamilial rivalries are also far from uncommon (I am a Liverpool fan, which antagonizes my father, a Manchester United fan). Using this model, I set out how supporting a football team creates a cultural memory influenced by individuals within that culture (the supporters), and how supporters are in turn influenced by being part of this cultural group.

The act of supporting your team is ongoing, and not restricted geographically or temporally, e.g., it is limited neither to the days matches are played, nor to the stadia, cities or even countries in which the matches take place. Supporting is a very social activity generally characterized by oral storytelling, semantic knowledge sharing, and reminiscence about events and characters associated with a club down the years. The matches themselves are, for the supporters, shared, repeated, very emotive experiences whether experienced in the stadium, in a local pub or half way around the world watching on television. The 'I was there' (or, for distant fans, 'I witnessed') motif of storytelling adds weight and authority, and interesting details

heard from one source may be incorporated into the next telling, adding to the cultural memory.

Songs and chants associated with the team are sung communally, and are often created by generally anonymous individual supporters. This is our first example of individuals having agency to affect a shared experience. It takes a single supporter to start a new song eulogising a player, and that song may be sung for decades; indeed many songs sung in football stadia have been around for generations. As a recent example, a pop group enlisted two celebrity supporters to sing a song in support of England at the European Championships in 1996. The lyrics drew heavily on cultural memories surrounding the English national team, and immediately engaged supporters.[10] The 'Three Lions' song is still sung, now by supporters who were not alive in 1996. The imagery and the words 'three lions' have become a central part of their shared memory.

There is a strong trans-generational aspect to this social support activity; contemporary supporters sing songs and tell stories of players and events they have never witnessed, or quote the wisdom of managers who died before they were born. Taking part in these activities does not require any prior links to any other participant. What is needed is (1) to self-identify as a supporter of that team, (2) to be identifiable by other participants as a supporter of that team, and (3) to share in the cultural memories of the team as embodied by the supporters.

Symbols, therefore, constitute a crucial part of the social experience. They not only allow members to identify each other, but can themselves form part of the shared cultural memory. While the business of the club includes selling merchandise such as team shirts, much of the most iconic symbolism is generated by supporters. Some of the most potent examples are flags and banners. For example, many created by Liverpool fans celebrating triumphs of the 1970s and 1980s are still displayed by supporters in the stadium and revered by their supporters elsewhere specifically because of the shared cultural memory for which they serve as reminders. New banners are created and displayed regularly, offering opportunity for any individual to create something which will become part of the evolving cultural memory of their club.[11]

Football supporters often define themselves, or at least an important part of their self, as a 'Liverpool fan' or a 'Chelsea fan'. Anyone ever asked by a Mancunian

---

10   The 'Three Lions' song is perhaps the most famous of English football songs, and has been updated by supporters over the last 25 years to maintain its relevance to contemporary tournaments. For full information see; https://en.wikipedia.org/wiki/Three_Lions

11   Fan-produced banners have become a large part of Liverpool supporters' collective identity and shared narrative, particularly since the 1960s. Supporters talk about banners being more for the 'collective'. Images of some Liverpool banners can be found here; https://lfc-fan-banners.com/ A fabulous documentary about Liverpool banners, with much input from makers, can be viewed here; https://youtu.be/e_z6CbrLOp8

Estud.filos n.º 64. Julio-diciembre de 2021 | pp. 47-71 | Universidad de Antioquia | ISSN 0121-3628 | ISSN-e 2256-358X   65

https://doi.org/10.17533/udea.ef.n64a03

the dreadful question 'City or United?' will feel the importance of their answer to the identity of the questioner. Are you one of us, with our shared history and understanding? Or are you one of them? One whose experiential memories – and identity – are in direct opposition to our own?

In these ways the culture is not only rich and vibrant for its members, but is also reliant on their participation in its generation and maintenance in order for it to continue as the same-but-not-the-same thing. By this I mean that the shared memory maintains a consistent yet constantly reconstructed composition by being comprised of shards derived from every member of the culture. It relies on distributed remembering and may be an example of an exceptionally widespread 'transactive memory system' (Heersmink & McCarroll, 2019, p. 99), where each individual remembers overlapping parts of the whole and each supports the memory of the other.

Looking at the model above we have a large group of geographically disparate people joined by a shared cultural memory powerful enough that these people consider it, to lesser or greater extents, part of their autobiographical selves or personal identities. The cultural memory is not static, but is maintained, renewed, expanded and transmitted across generations by the supporters themselves through participation in it. It is this requirement for participation which allows individuals to modify the cultural memory for all members. This model therefore appears to fully support the bidirectional relationship between individual and cultural memory hinted at by the senbazuru works. It offers multiple and ongoing opportunities for individuals to influence the nature of the shared memory, and the shared memory can become a significant aspect of the autobiographical 'self' of individuals within the collective.

Sociobiographical memory

I began by questioning the nature of photographs and how we perceive them, showing that traditional photographs are not memories. From trying to create a photograph which mimics aspects of memory came the metamorphogram; an object which takes a 'photograph' of its own experiences as a unique visual expression of 'self'. Using the metamorphogram to create many 'individuals' with shared experiences begged questions of shared memories and a stream of influence which flows both ways.

This long and interesting journey as an artist entering the philosophical world has exposed another bi-directional relationship. Interrogating photography with philosophical questions proved extremely valuable to the development of my artistic ideas. Interrogating philosophy through photographic art has led me to questions I had not previously considered. In the first part of this section I used art to develop a hypothesis around a form cultural memory which (1) can be directly affected by individuals, and (2) can affect individuals at the level of their autobiographical 'self'.

I then set out the Football Fan model, which appears to be a real-world example of this form of cultural memory. One may be able to think of many groups of people whose personal and shared identity is similarly entwined; punks (and steampunks), K-pop fans,[12] and MAGA enthusiasts (particular grass-roots supporters of Donald Trump) would fit equally well. In this final part I examine what sort of cultural memory are we seeing among these groups by comparison with existing definitions.

Empirical research shows that cultural context influences autobiographical memory. That is, self-identity is modified by shared cultural practices. Human groups seem predisposed to create social or cultural memories. Mnemonic practices, such as oral storytelling and behavioural habits, maintain and spread cultural memory across the generations. These mnemonic practices directly influence aspects of individual self-identity such as moral boundaries and taboos (for a review and examples see Alea & Wang, 2015; Wang, 2021; Zhang & Cross, 2011). However, in this sense cultural memory provides the context into which autobiographical memory is placed and reconstructed. Wang describes memory as being 'saturated in cultural context', with culture influencing what details are selected for remembering (Wang, 2021). 'Cultural memory' in this form may be thought of as a series of learned behaviours, morals and contexts; scaffolding onto which autobiographical memory is built. The shared experiential memory of a football fan would be constructed within such context, but is not this form of cultural memory.

In revisiting the works of Ricoeur, Suzi Adams (2019) describes Jan and Aleida Assmanns definition of Cultural Memory, which I capitalize to distinguish it from earlier uses. Adams calls Cultural Memory the 'institutionalized heritage of a society', the triad of remembered / forgotten / not forgotten (Assmann, 2006) information and artefacts which contribute a formative role in creating a 'collective identity' (Adams, 2019). The process of institutionalization – choosing which information is remembered – is generally carried out by a very small, select group, independent of input from almost all individuals within the society. The resultant collective identities appear prone to being, as Jan Assmann described all collective identities, 'products of the imagination', or abstracted stereotypes. They apply to large groups, nations or religions for example. Individuals are given little choice as to the nature of the collective identity of which they are a part, and the memories are not embodied within most people. Therefore, the Assmanns definition of institutional Cultural Memory does not fit our model.

---

12   K-pop (Korean pop music, exemplified by the band BTS) attracts very devoted followers with a predominantly younger demographic, who also share a self-created activist history. The fans famously made many hundreds of thousands of spoof applications for Trump rally tickets (https://www.theguardian.com/us-news/2020/jun/21/trump-tulsa-rally-scheme-k-popfans-tiktok-users), supported Black Lives Matter on social media by 'drowning out' hashtags of right-wing groups by posting K-pop memes (https://edition.cnn.com/2020/06/04/us/kpop-bts-blackpink-fans-black-lives-matter-trnd/index.html), and have history of social activism as a community (https://edition.cnn.com/2020/06/22/asia/k-pop-fandom-activism-intl-hnk/index.html).

Communicative memory, which Adams (2019) suggests is very similar to Halbwachs' 'collective memory', is described by Aleida Assmann as being about the everyday or the mundane, existing within living generations of people, often communicated orally. Retention beyond these limited temporal horizons requires it to become part of cultural memory. Assmann goes on to describe social memory as a part of communicative memory (Assmann, 2006). Social memory is about everyday life, intergenerational and based on lived experience. However, this sort of memory belongs to relatively small groups, such as families, work colleagues and so on. It addresses memories with social functions, providing contextual and socially useful information such as remembering jokes and the personal stories of others. Communicative memory, and subgenres within it, seem close to our football supporter model. However, communicative memory lacks a defined and consistent overall narrative shared by all members. Neither does communicative memory, being predominantly functional, appear to impact individuals at the level of their autobiographical self.

These existing definitions of cultural memory as a contextual scaffold, as institutional Cultural Memory, or various shades of smaller-scale communicative memory do not cover all social contexts. Whole areas of social interaction and lived experience encompass shared memories which fall outside of these systems, none of which offers agency to individuals with regard to shaping shared memory.

The shared memory exemplified by football supporters is not behavioural or generally unconscious as cultural contexts are described to be by Wang and others. It is neither tangential to nor abstracted from individual experience, in sharp contrast to institutionalized Cultural Memory. Instead it is direct and participatory, active and evolving. Strong autobiographical-yet-socially-shared memories are created and reinforced in specific social contexts. These memories are not semantic, having affective episodic content which is generally similar for all members. In this sense, there are strong similarities to the 'collective episodic memory' described by Manier & Hirst (2008, p. 257). Though developed entirely independently, both I and Manier & Hirst use football fans as an example. They note, for instance, that sports fans may form a collective memory of a specific event. However, the current model not only expands on this thinking, but differs from it significantly. The striking feature of this form of cultural memory —a narrative shared by all members of a group— is that it is malleable by individual effort. Its main defining characteristic may be its effect on individual self-identity, an effect produced by shared, direct and valued experiences. In light of the bi-directional influence between socially shared experience and autobiographical content, I propose 'sociobiographical memory' as a suitable descriptor.

Conclusion

This paper is more about a journey and a relationship between art and philosophy than it is about any philosophical conclusions. Having begun this paper examining our perception of photographs, how have I arrived at a point where I am discussing forms of cultural memory? The path began with John Sutton's *Philosophy and memory traces* (Sutton, 1998). When developing the metamorphogram my naïve, art-led interpretations no longer seemed appropriate to answer the questions arising from the process. What exactly were these images I was creating? And what exactly was I trying to represent? Sutton's writing showed the direction in which the answers might lie.

Through further reading of philosophical literature, I understood why traditional photographs are not memories, and why the metamorphogram might be. The explorations of philosophers, psychologists and sociologists allowed me to set definitions against which artistic ideas could be tested. They provided a structure supporting the artworks themselves. Most interestingly, the art developed in this context provided a visual language to communicate philosophical ideas to lay audiences outside of an academic context. Art driving engagement with philosophy, and philosophy informing art is, in my experience to date, a productive and generous two-way relationship.

In a world where photographs are everywhere, it can only be of benefit to understand our relationship with them. The concept of photographs as memories may be wrong, yet it is deeply embedded as a cultural metaphor and unlikely to change. However, Barthes' deeply personal reflections on photography have been built upon to show that we actually use photographs as tools ranging in power from mnemonic aids, to evocative objects which may become a physical part of an extended memory system (Heersmink, 2020). These developments in our understanding are already inspiring novel ways of helping people with memory problems.[13]

Photography, triggered by philosophical literature, then led me to explore cultural memory. Currently, 'sociobiographical memory' is an idea awaiting further development. However, the process of trying to understand what I was showing people (see Figure 4) begs questions of how philosophy relates to the real world. I found no extant definition for a myriad of everyday examples of shared memories which constitute part of an individual's autobiographical self. There exist mind-boggling numbers of self-creating, self-organizing groups, both short-lived and long-lived, which seem to fit preliminary criteria for sociobiographical memory. Modern technology means geographical limitations on forming social interactions with like-minded others, commonplace until the last decade or two, no longer exist. 'Furries', goths, Wall Street Bets[14] and other dynamic, almost

---

13   Studio Meineck is a design group focused on social benefit. Of particular relevance are projects for people living with dementia ('Music Memory Box' https://www.musicmemorybox.com/) and for narrative building ('Trove' https://studiomeineck.com/trove/).

14   WallStreetBets came to prominence for their collective attack on hedge-fund short selling of stocks (https://en.wikipedia.org/wiki/R/WallStreetBets). Their collective identity and shared pathway continued with sponsorship of 3500 gorillas; https://edition.cnn.com/2021/03/16/investing/wallstreetbets-gorillas-reddit/index.html.

exclusively online groups, and any number of comic and film-related fandoms are just some that spring immediately to mind. Their lived experiences and shared culture, and how these affect individual self, seems to be relatively unexamined philosophical territory, outside the domains of existing Cultural Memory.

In summary, communicating philosophical insights through art is one way to share these important discoveries about who we are with a much wider audience. Though only a sample of one, my practice has gained immeasurable benefit from working within philosophical contexts. From introducing concepts of memory into origami workshops, to better understanding the lived experience of people with dementia, philosophy adds something valuable to each conversation, and I am grateful to have stumbled into it.

References

Adams, S. (2019). A note on Ricœur's early notion of cultural memory. *Études Ricoeuriennes / Ricoeur Studies, 10*(1), 112-124. https://doi.org/10.5195/errs.2019.465

Alea, N. & Wang, Q. (2015). Going global: the functions of autobiographical memory in cultural context. *Memory, 23*(1),1-10. PMID: 25585205. https://doi.org/10.1080/09658211.2014.972416

Assmann, A. (2006). Memory, individual and collective. In R. E. Goodin & Ch. Tilly (Eds.), *The Oxford Handbook of Contextual Political Analysis* (pp. 210-224). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199270439.003.0011

Barthes, R. (1980) *Camera Lucida.* Vintage Books.

Baudrillard, J. (1998) *Paroxysm.* Verso, London.

Berger, J. (1980). *About looking.* Pantheon Books. https://doi.org/10.3817/1280046209

Fawns, T. (2020). Blended memory: a framework for understanding distributed autobiographical remembering with photography. *Memory Studies, 13*(6), 901-916. https://doi.org/10.1177/1750698019829891

Garry, M. & Gerrie, M. P. (2005). When photographs create false memories. *Current Directions in Psychological Science, 14*(6), 321–325. https://doi.org/10.1111/j.0963-7214.2005.00390.x

Gibson, W. (2012). *Distrust that particular flavour.* G. P. Puttnam's Sons.

Heersmink, R. (2020). Extended mind and artifactual autobiographical memory. *Mind and Language, 36,*1-15. https://doi.org/10.1111/mila.12353

Heersmink, R. & McCarroll, C. J. (2019). The best memories: identity, narrative, and objects. In T. Shanahan & P. Smart (Eds.), *Blade Runner 2049: a philosophical exploration* (pp. 87-107). Routledge. https://doi.org/10.4324/9780429460036-6

Hoskins, J. (2005). Agency, biography and objects. In C. Tiller, W. Keane, S. Küchler, M. Rowlands & P. Spyer (Eds.), *Handbook of Material Culture* (pp. 74-84). SAGE. https://doi.org/10.4135/9781848607972.n6

Manier, D. & Hirst, W. (2008). A cognitive taxonomy of collective materials. In A. Erll & A. Nünning (Eds.), *Cultural Memory Studies: An international and interdisciplinary handbook* (pp. 253-362). De Gruyter.

Rowlands, M. (2016). *Memory and the self: phenomenology, science and autobiography.* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780190241469.001.0001

St. Jacques, P. L. & Schacter, D. L. (2013). Modifying memory: selectively enhancing and updating personal memories for a museum tour by reactivating them. *Psychological Science, 24*(4), 537–543. https://doi.org/10.1177/0956797612457377

St Jacques, P. L., Montgomery, D. & Schacter, D. L. (2015). Modifying memory for a museum tour in older adults: reactivation-related updating that enhances and distorts memory is reduced in ageing. *Memory, 23*(6), 876–887. https://doi.org/10.1080/09658211.2014.933241

Sasaki, M & DiCicco, S. (2020). *The complete story of Sadako Sasaki and the thousand paper cranes.* Tuttle Publishing.

Sontag, S. (1979) *On photography* (electronic edition, 2005) Rosetta Books.

Sutton, J. (1998). *Philosophy and memory traces: Descartes to connectionism.* Cambridge University Press.

Wang, Q. (2021). The cultural foundation of human memory. *Annual Review of Psychology,* 72, 151–179. https://doi.org/10.1146/annurev-psych-070920-023638

Whittaker, S., Bergman, O. & Clough, P. (2010) Easy on that trigger dad: a study of long term family photo retrieval. *Pers Ubiquit Comput, 14*, 31–43. https://doi.org/10.1007/s00779-009-0218-7

Zhang, M. & Cross S. E. (2011). Emotions in memories of success and failure: a cultural perspective. *Emotion, 11*(4), 866-80. https://doi.org/10.1037/a0024025

Attitudes and the (dis)continuity between memory and imagination*

André Sant'Anna

Washington University in St. Louis, St. Louis, United States
E-mail: andrer@wustl.edu

Abstract: The current dispute between causalists and simulationists in philosophy of memory has led to opposing attempts to characterize the relationship between memory and imagination. In a recent overview of this debate, Perrin and Michaelian (2017) have suggested that the dispute over the (dis)continuity between memory and imagination boils down to the question of whether a causal connection to a past event is necessary for remembering. By developing an argument based on an analogy to perception, I argue that this dispute should instead be viewed as a dispute about the nature of the *attitudes* involved in remembering and imagining. The focus on attitudes, rather than on causal connections, suggests a new way of conceiving of the relationship between memory and imagination that has been overlooked in recent philosophy of memory.

Keywords: memory, imagination, causal theory, simulation theory, continuism, discontinuism

Cómo citar este artículo

Las actitudes y la (dis)continuidad entre la memoria y la imaginación

Resumen: La disputa actual entre causalistas y simulacionistas en filosofía de la memoria ha llevado a intentos opuestos de caracterizar la relación entre memoria e imaginación. En una revisión reciente de este debate, Perrin y Michaelian (2017) han sugerido que la disputa sobre la (dis)continuidad entre la memoria y la imaginación se reduce a la cuestión de si para recordar es necesaria una conexión causal con un evento pasado. Al desarrollar un argumento basado en una analogía con la percepción, sostengo que esta disputa debería verse como una disputa sobre la naturaleza de las actitudes involucradas en recordar e imaginar. El enfoque en las actitudes, más que en las conexiones causales, sugiere una nueva forma de concebir la relación entre la memoria y la imaginación que se ha pasado por alto en la filosofía reciente de la memoria.

Palabras clave: memoria, imaginación, teoría causal, teoría de la simulacion, continuismo, discontinuismo

André Sant'Anna

Postdoctoral Researcher in the Department of Philosophy and the Philosophy-Neuroscience-Psychology Program at Washington University in St. Louis.

ORCID: 0000-0002-2239-7243

Check for updates

1 Introduction

The current dispute between causalism and simulationism in philosophy of memory has led to opposing attempts to characterize the relationship between memory and imagination. According to causalists, memory is *discontinuous* with imagination, for a causal connection is necessary only for remembering. According to simulationists, memory is *continuous* with imagination, and for this reason, a causal connection is not necessary for remembering. This has led Perrin & Michaelian (2017) to suggest that the dispute over the (dis)continuity between memory and imagination boils down to the question of whether a causal connection is necessary for remembering. By developing an argument based on an analogy to perception, I propose that, given the commitment by causalists and simulationists to a representationalist approach to mental states, it is wrong to frame the dispute over the (dis)continuity between memory and imagination in terms of the necessity of a causal connection for remembering. Instead, I propose that it should be viewed as dispute about the nature of the *attitudes* involved in remembering and imagining. One crucial implication of this way of looking at things is, I will suggest, that philosophers of memory should distinguish between two related but separate debates: namely, the debate over whether a causal connection is necessary for remembering, on the one hand, and the debate over whether memory and imagination are continuous, on the other hand.

I proceed as follows: Section 2 introduces and discusses the causal theory, the simulation theory, and how they conceive of the relationship between memory and imagination. Section 3 draws an analogy to perception to argue that it is wrong to view the requirement for the presence of a causal connection in remembering as fundamental to establishing the (dis)continuity between memory and imagination. Section 4 discusses how my proposal relates to recent attempts to intervene in the (dis)continuism debate. Section 5 concludes by responding to potential objections to the analogy to perception argument.

2 (Dis)continuism and the necessity of a causal connection for remembering

Is a causal connection necessary for remembering? Two influential theories have been developed in response to this question. The causal theory of memory, or simply *causalism,* says that remembering occurs only when memory is appropriately caused by a past perceptual experience.[1] While there is room to dispute what it is for a past

---

event or experience to appropriately cause a current mental state, one popular strategy has been to appeal to the presence of a *memory trace*,[2] or a brain state that encodes and stores information at the time of experience and that is later retrieved to cause memories of those events. Thus, causalists have proposed that a causal connection is appropriate when it takes place by means of a memory trace connecting a particular past event to a current representation of it.

The causal theory has been dominant in philosophy for multiple reasons.[3] One of these reasons, which will be the focus of this paper, is that it captures an important metaphysical intuition about remembering: that is, that it differs in *kind* from imagining. A clear illustration is provided by the notorious painter example discussed by Martin & Deutscher (1966, pp. 167–168). In this example, we are asked to imagine the case of a painter who, as a result of being asked to paint an imagined scene, produces a painting of a farmyard that he genuinely believes to be imagined. However, when his parents see the painting, they recognize it as being a very accurate representation of a scene that the painter saw once as a child, thus suggesting that he is actually remembering that scene. The question that this example raises is whether the painter is remembering or imagining the scene. Setting aside the issue of whether the painter needs to believe that he is remembering in order to genuinely remember,[4] the causal theory offers a simple way to settle the issue: If there is an appropriate causal connection between the painter's current mental representation of the farmyard and his previous experience of it, the representation will count as case of remembering; in contrast, if such a causal connection is missing, the representation will count as a case of imagining. Thus, the presence of a causal connection serves to differentiate between memory and imagination, considered as *kinds* of mental states. In other words, it postulates a *discontinuity* between them.[5]

The dominant status of the causal theory has, however, been questioned recently. The second theory that attempts to answer the question of whether a causal connection is necessary for remembering, the simulation theory, or simply *simulationism,* proposes that remembering is just a form of *imagining* the past.[6] Recently developed in more detail by Michaelian (2016b), the simulation theory proposes that "[remembering] is not

---

2   Martin and Deutscher (1966) were the first to argue for this idea in the recent philosophy of memory literature. Despite the popularity of the causal theory, and despite being central for all subsequent versions of the theory (see Michaelian & Robins, 2018, for review), the idea of a memory trace has been the object of many criticisms. See Sutton (1998, ch. 16) for discussion; see also De Brigard (2014b); Robins (2017).

3   For a helpful discussion of the causal theory and its influence on subsequent philosophical theorizing, see Michaelian and Robins (2018).

4   See Debus (2010) and Fernández (2018) for arguments for the necessity of belief for remembering.

5   See Debus (2014); Perrin (2016); Michaelian (2016a) for discussion.

6   For the most influential exposition of the simulation theory, see Michaelian (2016b). For alternative formulations, see De Brigard (2014a) and Shanton and Goldman (2010). For my purposes, I will focus on Michaelian's version only. See also Hopkins (2018) for a view where remembering is viewed as a form of imagining, but that does is not committed to simulationism.

different in kind from other episodic constructive processes" (p. 103); thus "[w]hat it is for a subject to remember [...] is for him to imagine an episode belonging to his personal past" (p. 111). The motivation for the simulation theory comes from recent research on mental time travel.[7] According to this body of research, (episodic) memory and (episodic) imagination are just two specific occurrences of a more general cognitive capacity that we have for mental time travel in subjective time: while remembering is the specific ability we have to mentally travel into past subjective time, so as to "re-live" or "re-experience" an event, imagining corresponds to the specific ability we have to mentally travel into future subjective time, so as to simulate the experience of a possible event.[8]

The idea that remembering and imagining are two specific occurrences of a more general cognitive capacity for mental time travel has motivated further empirical studies on their relationship, which has reinforced the initial suggestion that the two are closely intertwined. Perhaps the most significant results come from neuroimaging studies, which have revealed a strong overlap of brain regions associated with memory and mental time travel into the future.[9] Building on this and other results, the simulation theory suggests that remembering and imagining are *continuous*; that is, that at the most fundamental level, they are mental states of the *same* kind.[10] The continuity between memory and imagination motivates the simulationist argument against the necessity of a causal connection for remembering. According to simulationists, given that mental time travel research implies that memory and imagination are mental states of the same kind, and given that imagination does not require a causal connection to what is imagined, it follows that a causal connection is not necessary for remembering.[11]

The current dispute between causalists and simulationists over the (dis)continuity between memory and imagination allows us to identify an important assumption made by both causalists and simulationists: namely, that if a causal connection is necessary for memory, then memory and imagination are mental states of different kinds.[12] Let us call this the *if-causation-then-discontinuity claim* (ICTD). ICTD has been at the basis of the disagreement between causalists and simulationists over the (dis)continuity of memory and imagination. On the one hand, causalists rely on ICTD to propose a *modus ponens* argument for discontinuism. Following the causal theory, they assert

---

7   See Perrin and Michaelian (2017) for a review. See also Sant'Anna *et. al.* (2020).

8   See Tulving (1993; 2002; 2005).

9   See Addis (2018; 2020); Addis et al. (2007); Schacter et al. (2007; 2012). See Perrin and Michaelian (2017) for a more detailed philosophical discussion.

10 See Michaelian (2016b;a); Michaelian *et al.* (2020), Sant'Anna (2020), Sant'Anna *et. al.* (2020).

11 The attempt to show that memory and imagination are continuous is not the only motivation that leads Michaelian to deny that a causal connection is necessary for remembering. Another equally important reason is the possibility of there being memory representations that are fully accurate but that are not causally connected to the original events—e.g. memories whose contents are derived from testimony or memories whose contents are derived from causal connections to events other than the event remembered. See Michaelian (2016b, ch. 6) for discussion.

12 See Perrin and Michaelian (2017) and Michaelian et. al. (2020) for explicit endorsements of this idea.

the necessity of a causal connection for memory. This claim, in conjunction with ICTD, allows for the conclusion that memory and imagination are mental states of different kinds. On the other hand, simulationists rely on ICTD to propose a *modus tollens* argument for continuism. Following the simulation theory, they note that it is not the case that memory and imagination are mental states of different kinds. Combined with ICTD, this claim allows for the conclusion that a causal connection is not necessary for remembering. The problem with how this debate is structured is, however, that no explicit argument has been given for ICTD. Rather, this is an assumption made by both causalists and simulationists. In the next section, I argue that ICTD is false, or that the dispute over the (dis)continuity between remembering and imagining should *not* be about the necessity of a causal connection for remembering.

3 Memory and imagination as representational states

Despite disagreeing about whether memory and imagination are (dis)continuous, causalists and simulationists alike are committed to the more general idea that they are *representational states*, or simply to *representationalism* about memory and imagination. In his main discussion of the simulation theory, Michaelian (2016b) speaks explicitly of memory and imagination as representational states. Similarly, the original version of the causal theory developed by Martin and Deutscher (1966) and subsequent developments of it clearly suggest a commitment to some form of representationalism about memory.[13] Furthermore, while philosophers of imagination have disputed the nature of the content of multiple forms of imagination,[14] most of them agree that different imaginative states are fundamentally representational states. As Liao & Gendler (2019) put it, to imagine "is to *represent* without aiming at things as they actually, presently, and subjectively are" (my emphasis; see also Currie & Ravenscroft, 2002). Representationalism is, to put it differently, a widespread assumption in both the memory literature and the imagination literature, with very few people questioning the view.[15]

   In what follows, I shall argue that, if representationalism is true, then the question about whether memory and imagination are (dis)continuous does not boil down to the necessity of a causal connection for remembering. Let me begin by defining representationalism more precisely. According to representationalism conceived in its most general form, what makes a mental state an occurrence of a certain kind is the *attitude* that one holds towards a *content* —e.g., a belief is characterized by an attitude of believing (understood as having a mind-to-world direction of fit) towards a certain

---

13   See Robins (2016) and Michaelian and Robins (2018) for discussion.

14   See, e.g., Nanay (2015); Langland-Hassan (2015).

15   See, however, Debus (2008); Hutto and Myin (2017).

content, a desire is characterized by an attitude of desiring (understood as having a world-to-mind direction of fit) towards a content, and so on (Fodor, 1978).[16] Thus, it follows that, if memory and imagination are understood as representational states, they too should be characterized in terms of the type of attitudes that is involved in each. The crucial question for any representationalist theory of memory and imagination is thus that of explaining the nature of these attitudes. Since, however, it is not my goal to defend representationalism, I shall leave this question aside.

What matters for my purposes is that the commitment to representationalism by both causalists and simulationists makes it clear that the question about the (dis)continuity between memory and imagination is not about the necessity of a causal connection for remembering, but rather about the nature of the attitudes involved in remembering and imagining. On the one hand, the causalist (and hence the discontinuist) will triumph if the attitude involved in remembering turns out to be *different* from the attitude involved in imagining. On the other hand, the simulationist (and hence the continuist) will triumph if the attitude involved in remembering turns out to be the *same* as the attitude involved in imagining. Whether or not a causal connection is necessary for remembering is orthogonal to settling this issue.[17]

To further motivate this point, considering an analogy to representationalist approaches to perception will help. One central question in recent philosophy of perception is whether veridical and non-veridical experiences are mental states of the same kind. Representationalists have answered this question in a positive manner.[18] According to them, because veridical and non-veridical experiences represent the world in the same way —that is, because they involve the same attitude towards contents— they are mental states of the *same* kind. They differ only in terms of whether their *contents* are satisfied. What it means to say that a content is satisfied by the world is a matter of controversy, but one natural way to understand this idea is to say that a particular object satisfies a perceptual experience, and hence makes it a veridical occurrence, when it causes the experience in an appropriate way. Thus, when I have a visual experience as of a cat across the street, this experience will be veridical *only if* it is caused by a cat that is across the street.[19] This allows representationalists to consistently hold on

---

16   It should be noted that the suggestion here is not, logically speaking, that we define the relevant attitudes in terms of the mental states they are intended to be a characterization of—e.g., that we define the attitude of believing in terms of what a belief is. In other words, representationalism is not engaged in circular reasoning here. Instead, to use belief as an example, the suggestion is that a belief is defined by an attitude of a certain type φ, which is characterized by such-and-such properties, and because φ is characteristic of mental states that we ordinarily refer to as beliefs, we refer to φ as the attitude of *believing*. Thanks to an anonymous referee for calling my attention to this issue.

17   One may object here that it is not true that, in the case of remembering in particular, the presence of a causal connection is not necessary to characterize the attitude of remembering. I respond to this objection in more detail in Section 5.2.

18   See, e.g., Bryne (2001; 2009); Schellenberg (2010); Searle (1983); Siegel (2010); Tye (2000).

19   A key issue here is whether the requirement for a causal connection to *individuate* a veridical experience as such —call this the *individuation claim*— requires that we *represent* the experience as being caused by the thing represented

to the claim that a causal connection is necessary, albeit not sufficient, for veridical experiences, while still maintaining that veridical and non-veridical experiences are mental states of the same kind. Otherwise put, the necessity of a causal connection for veridical experiences does not imply a *fundamental* separation between veridical and non-veridical experiences.

I want to suggest that a similar approach is available to understand the relationship between memory and imagination. That is, just like in perception, where the requirement for a causal connection for veridical experiences does not imply a fundamental separation between veridical and non-veridical experiences, in the case of memory and imagination too, the requirement for a causal connection for remembering does not imply a fundamental separation between memory and imagination. Once representationalism is accepted as a starting point, it becomes clear that there is no incompatibility between the two. This argument, which I call the *analogy to perception argument*, can be laid out as follows:

(The Analogy to Perception Argument)

(P1) If representationalism is true of memory and imagination, then the requirement for a causal connection for remembering, but not for imagining, only poses a non-fundamental difference between them.

(P2) Representationalism is true of memory and imagination.

(C) The requirement for a causal connection in remembering, but not in imagining, only poses a *non-fundamental* difference between them.

As I argued above, (P1) reflects a more general principle of representationalism, namely, that if two potentially distinct mental states —e.g., veridical and non-veridical experiences— involve the *same* attitude towards contents, then they are mental states of the same kind. That a causal connection is only necessary for one of them only reflects a non-fundamental difference. (P2), in contrast, reflects the widespread theoretical assumption made by most philosophers of memory and philosophers of imagination discussed previously.[20] Thus, once (P1) and (P2) are in place, it follows that (C) the requirement for a causal connection in remembering, but not in imagining, only poses a *non-fundamental* difference between them.

—call this the *representation claim*. While some, most notably Searle (1983), have answered this question positively, others have argued that the individuation claim can be secured without committing to the representation claim —see, e.g., Burge (1991); Soteriou (2000). For my purposes in this paper, I do not need to commit to any of these views.

20   I do not mean to suggest here that the widespread acceptance of (P2) by philosophers of memory and philosophers of imagination alike provides a *reason* for endorsing it. It may be, after all, that the assumption is wrong despite being widespread. The suggestion is, instead, that given its widespread acceptance, in particular by causalists and simulationists in the memory literature, and given the truth of (P1), (C) inevitably follows.

Despite establishing that the requirement for a causal connection for remembering does not imply a discontinuity between memory and imagination, it is important to note that the analogy to perception argument should be viewed neither as an argument for or against causalism, nor as an argument for or against simulationism. The argument is neutral as to whether a causal connection is necessary for remembering (causalism), as well as to whether memory and imagination are mental states of the same kind (simulationism). All it says is that, given a shared assumption between simulationists and causalists —namely, representationalism— the question of whether memory and imagination are mental states of the same kind should be a question about the nature of the *attitudes* involved in remembering and imagining. The analogy to perception thus renders false the ICTD claim introduced in Section 2.

Furthermore, a more general implication of the analogy to perception argument is that philosophers of memory should distinguish between two related but separate debates: namely, the debate over whether a causal connection is necessary for remembering, on the one hand, and the debate over whether memory and imagination are continuous, on the other hand. While, insofar as the current philosophy of memory literature is concerned, the latter has sprung out of the former, they concern different questions pertaining to the nature of remembering and imagining. The tendency to conflate them, exemplified by Perrin & Michaelian's (2017) proposal, overlooks important theoretical possibilities. In particular, once we distinguish between these two debates, a causalist-continuist view of memory becomes a real possibility. Whether such a view can be properly motivated is, of course, a question that is beyond my scope here, but it is certainly one that should be explored in future works on the subject.

4 Attitudes and the (dis)continuism debate

The suggestion that the (dis)continuism debate should be settled by considering the relationship between the attitudes of remembering and imagining has been echoed in recent work on the subject. For instance, Robins (2020) has recently argued that the attitude of 'seeming to remember', which she takes to be characteristic of occurrences of successful and unsuccessful remembering alike, and which involves entertaining a content as being past and as having happened, is clearly distinct from the attitude of imagining. The latter, she argues, involves entertaining a content as being fictional or possible (Van Leeuwen, 2013), thus suggesting that remembering and imagining are discontinuous. In a similar vein, Munro (2020) has argued that remembering is discontinuous with what he calls "hypothetical imagining" because they involve different attitudes towards contents. Unlike Robins (2020), though, Munro thinks that there is at least one type of imagining that is continuous with remembering —namely, what he calls 'actuality-oriented imagining', or situations in which one imagines actual

scenarios, such as imagining the layout of a restaurant where one is going to dine. Crucially, Munro's strategy for defending this view is that of showing that remembering and actuality-oriented imagining involve attitudes of a very similar type. Thus, despite their differences, these two attempts share a more general motivation to resolve the (dis)continuism debate by offering characterizations of the attitudes of remembering and imagining, and as such, they come in support of the claim defended in this paper.

One dissenting proposal has, however, been advanced by Langland-Hassan (2021), which might cast doubt on the main claim I am defending here. According to Langland-Hassan, it is wrong to view the (dis)continuism debate as a debate about attitudes. Instead, he argues that it should be viewed as a controversy over whether remembering is an instance of what has been called 'constructive imagining' (Van Leeuwen, 2013). The reason we should refrain from talking about attitudes, Langland-Hassan adds, is that continuists (and here he has the simulationist in mind) will happily accept the claim that memory and imagination clearly involve different attitudes. The argument in support of this claim appeals to Michaelian's (2016) claim that one of the conditions for remembering to happen is that it is produced by a reliably functioning episodic construction system that 'aims' at representing an event from one's personal past. This condition, Langland-Hassan argues, places unique epistemic constraints on remembering that do not hold for imagining. As he puts it, "[t]o say that the episodic construction system "aims at" an episode from one's actual personal past is to say that its products are in epistemic need of revision when that aim isn't met −viz., when the episodic memory does not accurately represent an episode from one's actual personal past" (Langland-Hassan, 2021, p. 237). Thus, since the same is not true of imagination, it follows that the attitudes of remembering and imagining are of different types.

As it stands, there are at least two difficulties with this argument. A first difficulty is that the simulationist view, at least as formulated by Michaelian (2016), does not require that the episodic construction system be successful in achieving its goal of representing an event from one's personal past for its outputs to count as occurrences of remembering. All that needs to be the case is, first, that the episodic construction system is functioning reliably and, second, that it has the 'aim' of representing an event from one's personal past, although it may well fail to do so. In other words, the requirement is that the episodic construction system has the relevant aim, and not that it succeeds in achieving that aim. That such is the case becomes clear when we consider the fact that Michaelian (2016, pp. 68-70) outright rejects a factive conception of remembering, which he takes to be incompatible with the naturalistic outlook on which the simulation view is based.

The reason this creates a problem for Langland-Hassan's proposal is that a similar way of talking of 'aims' could be proposed in an attempt to further specify when the episodic construction system is engaged in representing hypothetical and/ or future events. Consider future-oriented episodic imagining. It could be argued that one successfully imagines when the relevant representation is produced by a

reliably functioning episodic construction system and when the system has the aim of representing an event in one's possible personal future. Likewise, consider past-oriented counterfactual imagining. It could be argued that one successfully imagines a past counterfactual scenario when the relevant representation is produced by a reliably functioning episodic construction system and when the system has the aim of representing an event in one's counterfactual personal past. In both cases, what matters is, just like in the case of remembering, that the system succeeds in having the relevant aim, and not that it succeeds in actually representing possible future or past counterfactual events. Now, once we interpret the requirement in question in this way, it is no longer clear whether, for the simulationist, the epistemic constraints placed on remembering differ in nature from the epistemic constraints placed on imagining. For what it takes for the system to succeed in all those cases is simply for it to have the goal of representing events as being a certain way.

The second, and related, difficulty faced by Langland-Hassan's proposal is that not only is it not clear that simulationists will readily endorse the idea that memory and imagining involve attitudes of different types, it is also not clear whether causalists would take that as a starting point. To see the point, consider again the painter case discussed above. If it is true that causalists readily accept that the attitude of remembering is different from the attitude of imagining, then the issue over whether the painter is remembering would easily be settled against them, for the painter is clearly entertaining a content as possible or fictional. Thus, it is hard to see how there would even be a question as to whether the painter is remembering if causalists took the attitude of remembering to be clearly distinct from the attitude of imagining. So, it may be that not even causalists would be convinced, at least prima facie, by the thought that the attitudes of remembering and imagining are clearly different. This, I submit, shows that the idea that the (dis)continuism debate has to do with what the attitudes of remembering and imagining are cannot be as easily dismissed as Langland-Hassan (2021) suggests.

5 Objections

In this final section, I will consider a few potential objections to the analogy to perception argument and argue that none of them are successful.

5.1 Representationalism and discontinuism

A first objection is that if the analogy to perception argument is right, then, given that it is obviously true that remembering and imagining involve *different* attitudes towards contents, in the same way that it is obviously true that remembering and perceiving

involve *different* attitudes towards contents, it follows that discontinuism is the case. And this undermines the intended neutrality of the argument, for it is best viewed as an argument for causalism or discontinuism.

To see why this is not the case, consider the painter case discussed in Section 2. The fact that we are willing to accept that the painter may be remembering despite taking himself to be imagining a farmyard scene suggests that at least some occurrences of remembering play the same cognitive role as occurrences of imagining. Thus, as long as their contents are the same, it follows that they involve the *same* attitudes.[21] Admittedly, this only establishes that *some* occurrences of imagining and some occurrences of remembering involve the same attitudes towards contents. One may argue, however, that there are occurrences of imagining that clearly play distinctive cognitive roles from occurrences of remembering. While this is true, it does not follow that the dispute is settled in favor of discontinuism. For these differences could, at least in principle, be explained in terms of remembering and imagining typically (although not necessarily) involving different contents —e.g., remembered events are typically represented as *actual* and imagined events are typically represented as *possible.* This would still be compatible with the idea that remembering and imagining involve the same attitudes, and hence compatible with continuism.

One clarification here is that I am not claiming that causalists themselves think that remembering and imagining are attitudes of the same kind. The claim is rather that, as long as we are speaking of attitudes, making sense of the painter case in the way that it has been usually conceived of in recent discussions requires acknowledging that some occurrences of remembering can involve the attitude of imagining. Otherwise, it is not clear why we should say that the painter remembers despite his mental state clearly involving an attitude of imagining —that is, despite him taking himself to be imagining the farmyard scene.[22]

To further motivate this point, consider an analogy to beliefs and desires. If we define beliefs and desires in terms of their attitudes, then it would be odd to say that it is possible for one to believe that p, but mistakenly take oneself to desire that p. If, in the case in question, p plays the cognitive role of desiring —i.e., one expects p to obtain, etc.— then it looks like one has a mental state of desiring that p and not one of believing that p. On an attitude view of mental states, it would only be possible for one to believe that p but mistakenly take oneself to desire that p if the attitude of believing and the attitude of desiring were the same. The argument I am putting forward in connection

---

21   One may object here by saying that, if representationalism is true, then it cannot be the case that the painter is imagining, for he holds an attitude of remembering towards a content. In response, this only follows if discontinuism is accepted as a starting point, that is, if the attitude of remembering and the attitude of imagining are taken to be fundamentally different. However, on a continuist framework, this would not be problematic, for the attitude of remembering is just the attitude of imagining. Thus, it is not the case that representationalism necessarily conflicts with our intuitions in those cases.

22   Thanks to an anonymous referee for suggesting that I clarify this point.

to memory and imagination is similar: on an attitude view of mental states, it would only be possible for one to remember that p but mistakenly take oneself to imagine that p –this is what happens in the painter case– if the attitude of remembering and the attitude of imagining were the same. So, again, conceiving of the dispute between continuism and discontinuism in terms of attitudes does not straightforwardly settle the debate in favor of discontinuism.

5.2 Attitudes and causal connections

A second objection is that the analogy to perception argument presupposes that we can engage in discussions about the differences between the attitudes involved in remembering and imagining without talking about causal connections. However, the objection goes, the requirement for a causal connection is *essential* to characterizing the attitude of remembering, for part of what it means to say that a memory is accurate is that it is caused in an appropriate way by the event represented. So, even within a representationalist framework, the (dis)continuism debate ultimately boils down to the question of whether a causal connection is necessary for remembering.

This objection can be avoided by pointing out that, while it might be right in that reference to a causal connection is required to account for the accuracy conditions of memory,[23] it is wrong in that such a reference is essential to characterize the attitude of remembering. In Section 3, I suggested that a causal connection is required to determine when the *content* of a memory representation is satisfied (see also Fernández, 2019). So, the requirement is not built into the attitude, but the content, of remembering. We can, therefore, make sense of the idea that reference to a causal connection is required to speak of the accuracy conditions of memory, without making that a fundamental aspect of the attitude of remembering.

It may be argued in response that, even if the argument developed in the paper does not require building a causal connection into the attitude of remembering, a view along these lines is not incoherent. So, one may be a representationalist and still maintain that the attitude of remembering is fundamentally different from the attitude of imagining because only the former requires a causal connection. In response, I want to acknowledge that this view is not logically incoherent and that it may provide a way for causalists and simulationists to avoid the analogy to perception argument. And indeed, a proposal along these lines has recently been advanced by Mahr & Csibra (2018) and Mahr (2020).[24] For proposals of this type to be successful, they need an argument showing why the requirement for a causal connection should be built into the attitude,

---

23   Although the simulationist will, of course, disagree with this claim.

24   For similar proposals, although they do not speak explicitly of attitudes, see Dokic (2001; 2014), Perner & Ruffman (1995), Perrin et al. (2020).

as opposed to the content, of remembering.[25] This is, however, a controversial issue (*Cf.* Fernández, 2019) that I do not intend to resolve here. So, for my current purposes, I shall simply note that the question of whether reframing the debate over the (dis) continuity between memory and imagination in terms of attitudes will inevitably lead to discontinuism depends on resolving the more basic issue of whether we should build the requirement for a causal connection for remembering in its attitude or its content.

5.3 The factivity of remembering

A third objection appeals to the idea that remembering involves a factive attitude. Since a factive attitude requires its content to be accurate, it could be argued that a causal connection is essential to characterize the attitude of remembering. This would imply that one may be a representationalist and still maintain that the attitude of remembering is fundamentally different from the attitude of imagining —which is non-factive— because only the former requires a causal connection.

While the issue of whether (episodic) remembering is factive is controversial, even if we accept that it is, it does not follow that a causal connection is essential to characterize the attitude of remembering. This objection assumes that a factive attitude requires the presence of a causal connection, but it is not clear why we should make this assumption. To see the point, consider an analogy to semantic or propositional memory, which is thought to involve a factive attitude (Fraise, 2015). It does not follow from this that characterizing the attitude of semantic remembering requires appealing to a causal connection. I can semantically remember that Paris is the capital of France even if that memory is not caused by the fact that Paris is the capital of France. All that is required is that this fact obtains in the world. A similar analogy can be made to knowledge. That knowledge involves a factive attitude does not imply that to know something, one needs to be causally connected in an appropriate way to what is known. In other words, one does not need to subscribe to a causal theory of knowledge to hold the view that knowledge is factive. Thus, even if remembering is taken to be factive, it does not follow that a causal connection is essential to characterize the attitude involved in it.

It may be replied that, even if the above is true, the notion that remembering is factive and imagining is not points to a fundamental difference between them. This brings us back to the first objection discussed above. If the analogy to perception argument is correct, then it follows that discontinuism is true, which threatens the neutrality of the argument. Two things can be said in response. First, as noted above, it is not uncontroversial that remembering is factive. As different authors have pointed out,

---

25   Mahr & Csibra (2018, p. 3; see also Mahr, 2020) offer such an argument.

there are good reasons to think that it is not.[26] So, the truth of discontinuism ultimately depends on it being the case that remembering is factive. Second, that the factivity of remembering and the non-factivity of imagining is taken to be a potential reason for endorsing discontinuism only reinforces the conclusion of the analogy to perception argument. The issue of whether remembering and imagining are factive has to do with the nature of the attitudes involved in those mental states, and not with whether they require a causal connection. As noted in Section 3, the analogy to perception argument is neither an argument for nor against continuism and discontinuism. It establishes only that this dispute should be viewed as being about the nature of the attitudes involved in remembering and imagining. The fact that the factivity of remembering is said to support discontinuism only reinforces this idea.

5.4 Representationalism without attitudes

A fourth objection is that committing to the idea that memory and imagination are representational states does not require causalists and simulationists to endorse representationalism in the way that it is defined here. This point can be made in connection to the perception literature. While some have indeed defended the view according to which perception is a propositional attitude, others have explicitly denied this view in a representationalist framework.[27] On views of this type, which we may call *content representationalism*, what characterizes a mental state as a perceptual state is the nature of its content, understood as its accuracy conditions. Likewise, one may argue that what characterizes a mental state as a memory/imagination is the nature of its content, understood as its accuracy conditions. Thus, if a causal connection is built into the content of memory and not into the content of imagination, it follows that there is a fundamental difference between them, which ultimately boils down to the presence (or absence) of a causal connection for remembering.

Endorsing content representationalism would allow causalists and simulationists to avoid the analogy to perception argument. The question is whether there are good reasons for them to do so. Let us start with causalists. As the painter case discussed previously illustrates, one important idea for causalism is that genuine remembering and apparent remembering, which includes merely imagining the past, can and often are phenomenologically indistinguishable. On a representationalist framework, where phenomenology supervenes on content, this would mean that some occurrences of imagining the past and some occurrences of remembering the past have contents of the same type, thus implying that they are mental state of the same kind. However, given

---

26   See Michaelian (2016b); De Brigard (2014a; 2017); Hazlett (2010).

27   See, e.g., Crane (2009); Siegel (2010); Schellenberg (2018).

that causalists ultimately want to endorse discontinuism, this result is undesirable. So, causalism is not compatible with content representationalism.

Consider now simulationists. For continuism to follow from content representationalism, it must be the case that memory and imagination have contents of the same type. It is not clear, however, what the motivations for endorsing this claim are. The simulationist seems to agree that, despite being continuous, memory and imagination can represent things differently. For instance, the simulationist does not deny that sometimes memory represents events as being past and that sometimes imagination represents events as being future. But if that is the case, then there is a difference in the nature of the content of memory and imagination —namely, they represent time in a different way. So, unless simulationists can offer an account of the nature of the contents of memory and imagination that avoids these problems, which it is not clear they can,[28] content representationalism is unlikely to be attractive for them.

In sum, it looks like that, given their commitments elsewhere, content representationalism is not a viable alternative for causalists and simulationists to resist the analogy to perception argument.

5.5 Rejecting representationalism

A fifth and final objection would be to deny representationalism. In doing so, one can reject (P2) from the analogy to perception argument, and as such, reject its conclusion. In response, it is fair to say that this an open possibility, and nothing in the discussion above establishes that (P2) is actually the case. However, as noted before, representationalism is widely accepted, particularly in the memory literature, so, while rejecting (P2) would be plausible, it would require substantial argumentation. Thus, the analogy to perception argument can be viewed as posing a dilemma for causalists and simulationists. That is, *either* it is true that memory and imagination are representational states, in which case the dispute over their (dis)continuity is *not* about the necessity of a causal connection for remembering, *or* it is the case that the dispute over (dis)continuity between memory and imagination *is* about the necessity of a causal connection for remembering, in which case it cannot be the case that memory and imagination are representational states. It is up to the causalist and the simulationist to choose which horn of the dilemma to endorse, and it may be that they will decide to endorse the second horn; however, should they do so, they will be faced with the difficult task of providing a non-representational account of memory and imagination.

---

28 See, e.g., De Brigard & Gessell (2016) and Mahr (2020) for recent arguments that the contents of episodic representations are tenseless.

6 Conclusion

In this paper, I argued that causalists and simulationists wrongly identify the issue of whether a causal connection is necessary for remembering as being central to their dispute over whether memory and imagination are (dis)continuous. I showed that, given their commitment to a representationalist approach to mental states, the question of whether memory and imagination are (dis)continuous should be viewed as a question about whether they involve the same or different attitudes towards contents. To secure this claim, I developed an argument in analogy to perception —what I called the *analogy to perception argument*— according to which, like in the case of perception, where given the truth of representationalism, requiring the presence of a causal connection for veridical experiences, but not for non-veridical experiences, only poses a non-fundamental difference between them, in the case of memory and imagination too, given the truth of representationalism, requiring the presence of a causal connection for remembering, but not for imagining, only poses a non-fundamental difference between. While the analogy to perception argument does not provide an answer to the question of whether memory and imagination are (dis)continuous, it establishes that, as long as representationalism is taken to be a starting point in this discussion, for there to be a fundamental difference between memory and imagination, it is required that the attitudes involved in each are different.

References

Addis, D. R. (2018). Are episodic memories special? on the sameness of remembered and imagined event simulation. *Journal of the Royal Society of New Zealand, 48*(2-3): 64–88. https://doi.org/10.1080/03036758.2018.1439071

Addis, D. R. (2020). Mental time travel? A neurocognitive model of event simulation. *Review of Philosophy and Psychology, 11*, 233–259. https://doi.org/10.1007/s13164-020-00470-0

Addis, D. R., Wong, A. T. & Schacter, D. L. (2007). Remembering the past and imagining the future: common and distinct neural substrates during event construction and elaboration. *Neuropsychologia, 45*(7), 1363–1377. https://doi.org/10.1016/j.neuropsychologia.2006.10.016

Bernecker, S. (2010). *Memory: a philosophical study.* Oxford University Press.

Burge, T. (1991). Vision and intentional content. In E. Lepore & R. van Gulick (Eds.), *John Searle and his critics* (pp. 195–214). Blackwell.

Byrne, A. (2001). Intentionalism defended. *Philosophical Review, 110*(2), 199–240. https://doi.org/10.1215/00318108-110-2-199

Estud.filos n.º 64. Julio-diciembre de 2021 | pp. 73-93 | Universidad de Antioquia | ISSN 0121-3628 | ISSN-e 2256-358X

89

https://doi.org/10.17533/udea.ef.n64a04

Byrne, A. (2009). Experience and content. *The Philosophical Quarterly, 59*(236), 429–451. https://doi.org/10.1111/j.1467-9213.2009.614.x

Crane, T. (2009). Is perception a propositional attitude? *The Philosophical Quarterly, 59*(236), 452–469. https://doi.org/10.1111/j.1467-9213.2008.608.x

Currie, G. & Ravenscroft, I. (2002). *Recreative minds: imagination in philosophy and psychology*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198238089.001.0001

De Brigard, F. (2014a). Is memory for remembering? Recollection as a form of episodic hypothetical thinking. *Synthese, 191*(2), 155–185. https://doi.org/10.1007/s11229-013-0247-7

De Brigard, F. (2014b). The nature of memory traces. *Philosophy Compass, 9*(6), 402–414. https://doi.org/10.1111/phc3.12133

De Brigard, F. (2017). Memory and imagination. In S. Bernecker & K. Michaelian (Eds.), *The Routledge Handbook of Philosophy of Memory* (pp. 127–140). Routledge. https://doi.org/10.4324/9781315687315-11

De Brigard, F. & Gessell, B. (2016). Time is not of the essence: Understanding the neural correlates of mental time travel. In K. Michaelian, S. B. Klein & K. K. Szpunar (Eds.), *Seeing the future: theoretical perspectives on future-oriented mental time travel* (pp. 153–179). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780190241537.003.0008

Debus, D. (2008). Experiencing the past: a relational account of recollective memory. *Dialectica, 62*(4). 405–432. https://doi.org/10.1111/j.1746-8361.2008.01165.x

Debus, D. (2010). Accounting for epistemic relevance: a new problem for the causal theory of memory. *American Philosophical Quarterly, 47*(1), 17–29.

Debus, D. (2014). 'Mental time travel': remembering the past, imagining the future, and the particularity of events. *Review of Philosophy and Psychology, 5*(3), 333–350. https://doi.org/10.1007/s13164-014-0182-7

Dokic, J. (2001). Is memory purely preservative? In C. Hoerl & T. McCormack (Eds.), *Time and memory: issues in philosophy and psychology* (pp. 213–232). Oxford University Press.

Dokic, J. (2014). Feeling the past: a two-tiered account of episodic memory. *Review of Philosophy and Psychology, 5*(3), 413–426. https://doi.org/10.1007/s13164-014-0183-6

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. (1978). Propositional attitudes. *The Monist, 61*(4), 501-523. https://doi.org/10.5840/monist197861444

Frise, M. (2015). Epistemology of memory. *The Internet Encyclopedia of Philosophy*. https://iep.utm.edu/

Hazlett, A. (2010). The myth of factive verbs. *Philosophy and Phenomenological Research, 80*(3),497–522. https://doi.org/10.1111/j.1933-1592.2010.00338.x

Hopkins, R. (2018). Imagining the past: on the nature of episodic memory. In F. MacPherson & F. Dorsch (Ed.), *Perceptual Imagination and Perceptual Memory*. Oxford University Press.

Hutto, D. D. & Myin, E. (2017). *Evolving enactivism: basic minds meet content*. MIT Press. https://doi.org/10.7551/mitpress/9780262036115.001.0001

Langland-Hassan, P. (2015). Imaginative attitudes. *Philosophy and Phenomenological Research, 90*(3), 664–686. https://doi.org/10.1111/phpr.12115

Langland-Hassan, P. (2021). What sort of imagining might remembering be? *Journal of the American Philosophical Association, 7(2), 231-51.*

Liao, S.-Y. & Gendler, T. (2019). Imagination. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University.

Mahr, J. B. (2020). The dimensions of episodic simulation. *Cognition, 196*, 104085. https://doi.org/10.1016/j.cognition.2019.104085

Mahr, J. B. & Csibra, G. (2018). Why do we remember? The communicative function of episodic memory. *Behavioral and Brain Sciences, 41.* https://doi.org/10.1017/S0140525X17000012

Martin, C. B. & Deutscher, M. (1966). Remembering. *Philosophical Review, 75*(2), 161–96. 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. (2016a). Against discontinuism: mental time travel and our knowledge of past and future events. In K. Michaelian, S. B. Klein & K. K. Szpunar (Eds.), *Seeing the future: theoretical perspectives on future-oriented mental time travel* (pp. 62–92). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780190241537.003.0004

Michaelian, K. (2016b). *Mental time travel: episodic memory and our knowledge of the personal past*. MIT Press. https://doi.org/10.7551/mitpress/10591.001.0001

Michaelian, K., Perrin, D. & Sant'Anna, A. (2020). Continuities and discontinuities between imagination and memory: the view from philosophy. In A. Abraham (Ed.), *The Cambridge handbook of imagination*. Cambridge University Press. https://doi.org/10.1017/9781108580298.019

Michaelian, K. & Robins, S. K. (2018). Beyond the causal theory? Fifty years after Martin and Deutscher. In K. Michaelian, D. Debus, & D. Perrin (Ed.), *New directions in the philosophy of memory* (pp. 13–32). Routledge. https://doi.org/10.4324/9781315159591-2

Munro, D. (2020). Remembering the past and imagining the actual. *Review of Philosophy and Psychology*. Online ahead of print. https://doi.org/10.1007/s13164-020-00499-1

Nanay, B. (2015). Perceptual content and the content of mental imagery. *Philosophical Studies, 172*(7), 1723–1736. https://doi.org/10.1007/s11098-014-0392-y

Perner, J. & Ruffman, T. (1995). Episodic memory and autonoetic aonciousness: developmental evidence and a theory of childhood amnesia. *Journal of Experimental Child Psychology, 59,* 516–548. https://doi.org/10.1006/jecp.1995.1024

Perrin, D. (2016). Asymmetries in subjective time. In K. Michaelian, S. B. Klein & K. K Szpunar, (Eds.), *Seeing the future: theoretical perspectives on future-oriented mental time travel* (pp. 39–61). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780190241537.003.0003

Perrin, D. & Michaelian, K. (2017). Memory as mental time travel. In S. Bernecker & K. Michaelian (Eds.), *The Routledge handbook of philosophy of memory* (pp. 228–239). Routledge. https://doi.org/10.4324/9781315687315-19

Perrin, D., Michaelian, K. & Sant'Anna, A. (2020). The phenomenology of remembering is an epistemic feeling. *Frontiers in Psychology, 11,* 1531. https://doi.org/10.3389/fpsyg.2020.01531

Robins, S. K. (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

Robins, S. K. (2017). Memory traces. In S. Bernecker & K. Michaelian (Eds.), *The Routledge handbook of philosophy of memory* (pp. 76–87). Routledge. https://doi.org/10.4324/9781315687315-7

Robins, S. K. (2020). Defending discontinuism, naturally. *Review of Philosophy and Psychology, 11,* 469–486. https://doi.org/10.1007/s13164-020-00462-0

Sant'Anna, A. (2020). The hybrid contents of memory. *Synthese 197,* 1263–1290. https://doi.org/10.1007/s11229-018-1753-4

Sant'Anna, A., Michaelian, D. & Perrin, D. (2020). Editorial: memory as mental time travel. *Review of Philosophy and Psychology, 11,* 223–232. https://doi.org/10.1007/s13164-020-00484-8

Schacter, D. L., Addis, D. R. & Buckner, R. L. (2007). Remembering the past to imagine the future: the prospective brain. *Nature Reviews Neuroscience, 8*(9)*,* 657–661. https://doi.org/10.1038/nrn2213

Schacter, D. L., Addis, D. R., Hassabis, D., Martin, V. C., Spreng, R. N. & Szpunar, K. K. (2012). The future of memory: remembering, imagining, and the brain. *Neuron, 76*(4), 677–694. https://doi.org/10.1016/j.neuron.2012.11.001

Schellenberg, S. (2010). The particularity and phenomenology of perceptual experience. *Philosophical Studies, 149*(1), 19–48. https://doi.org/10.1007/s11098-010-9540-1

Schellenberg, S. (2018). *The unity of perception: content, consciousness, evidence.* Oxford University Press. https://doi.org/10.1093/oso/9780198827702.001.0001

Searle, J. R. (1983). *Intentionality: an essay in the philosophy of mind.* Cambridge University Press. https://doi.org/10.1017/CBO9781139173452

Shanton, K. & Goldman, A. (2010). Simulation theory. *Wiley Interdisciplinary Reviews, Cognitive Science, 1*(4), 527–538. https://doi.org/10.1002/wcs.33

Siegel, S. (2010). *The contents of visual experiences.* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195305296.003.0002

Soteriou, M. (2000). The particularity of visual experience. *European Journal of Philosophy, 8*(2), 173–89. https://doi.org/10.1111/1468-0378.00107

Sutton, J. (1998). *Philosophy and memory traces: Descartes to connectionism.* Cambridge University Press. https://doi.org/10.1111/1467-8721.ep10770899

Tulving, E. (1993). What is episodic memory? *Current Directions in Psychological Science, 2*(3), 67–70. https://doi.org/10.1146/annurev.psych.53.100901.135114

Tulving, E. (2002). Episodic memory: from mind to brain. *Annual Review of Psychology, 53*(1), 1–25. https://doi.org/10.1093/acprof:oso/9780195161564.003.0001

Tulving, E. (2005). Episodic memory and autonoesis: uniquely human? In H. S. Terrace & J. Metcalfe (Eds.), *The missing link in cognition: origins of self-re*flective *consciousness* (pp. 3–56). Oxford University Press. https://doi.org/10.7551/mitpress/2110.001.0001

Tye, M. (2000). *Consciousness, color, and content.* MIT Pres.

Van Leeuwen, N. (2013). The meanings of "imagine", part 1: constructive imagining. *Philosophy Compass, 8,* 220–230. https://doi.org/10.1111/j.1747-9991.2012.00508.x

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

What makes a mental state feel like a memory: feelings of pastness and presence*

Melanie G. Rosen

Trent University, Peterborough, Canada
E-mail: melanierosen@trentu.ca

Michael Barkasi

York University, Toronto, Canada
E-mail: michael.barkasi@gmail.com

Abstract: The intuitive view that memories are characterized by a feeling of pastness, perceptions by a feeling of presence, while imagination lacks either faces challenges from two sides. Some researchers complain that the "feeling of pastness" is either unclear, irrelevant or isn't a real feature. Others point out that there are cases of memory without the feeling of pastness, perception without presence, and other cross-cutting cases. Here we argue that the feeling of pastness is indeed a real, useful feature, and although this feeling does not define memory ontologically, it is a characteristic marker which helps us easily categorise a mental state first-personally. We outline several cognitive features that underlie this experience, including the feeling of past accessibility, ergonomic significance, immersion, objectivity and mental strength. Our account is distinctly phenomenal, rather than doxastic, although our web of beliefs may contribute to this experience.

Key words: memory, perception, imagination, phenomenology

Cómo citar este artículo

OPEN ACCESS

Qué hace que un estado mental se sienta como un recuerdo: sentimientos de pasado y presencia

Resumen: La visión intuitiva de que los recuerdos se caracterizan por un sentimiento de pasado, las percepciones por un sentimiento de presencia, mientras que la imaginación carece de cualquiera de los dos, enfrenta varios desafíos. Algunos investigadores se quejan de que el "sentimiento de pasado" no es claro, es irrelevante o no es una característica real. Otros señalan que hay casos de memoria sin sentimiento de pasado, percepción sin sentimiento de presencia y otros casos transversales. Aquí sostenemos que, aunque el sentimiento de pasado no define ontológicamente la memoria, este es de hecho una característica real y útil y, además, es un marcador característico que nos ayuda a categorizar fácilmente un estado mental. Describimos varias características cognitivas que subyacen a esta experiencia, incluida la sensación de accesibilidad pasada, el significado ergónomico, la inmersión, la objetividad y la fuerza mental. Nuestra perspectiva es claramente fenoménica, más que doxástica, aunque nuestra red de creencias puede contribuir a esta experiencia.

Palabras clave: memoria, percepción, imaginación, fenomenología

Melanie G. Rosen is an assistant professor at Trent University, formerly a Carlsberg distinguished postdoctoral research fellow at Aarhus University. Her research takes an interdisciplinary, philosophy focused approach to altered states of consciousness, perception and memory.

ORCID: 0000-0002-7545-5558

Michael Barkasi is a philosopher of perception who works at the intersection of consciousness, dreams, hallucinations, memory, and neural encoding. He is a former postdoctoral research fellow at the Network for Sensory Research (University of Toronto). Most recently, he was an instructor of philosophy and cognitive science at York University in Toronto.

ORCID: 0000-0003-4070-8196

Check for updates

1 Introduction

A sense of presence and pastness, we argue, are key features that, from a first-person perspective, allow individuals to categorise perception, memory and imagination. Here we argue that sense of presence and pastness should be seen as important phenomenal features of our experience rather than what categorise mental states objectively. While several cognitive features, such as how vivid an experience is, have been argued as the main factors that distinguish our mental states, we assess several of them and conclude that it is likely that none are necessary or sufficient. Rather, a cluster of indicators contributes to the sense of pastness for memory, presence for perception and absence of these features for imagination. We discuss how these contributions come together to form our experience.

One might consider that, from the first-person perspective, memory and imagination are very similar. As Debus (2016, p. 136) says, "considered in isolation and from the subject's own point of view, [memories and imaginings] might be difficult to tell apart." In contrast, some consider memory to be phenomenally like perception. Teroni (2017, p. 23) describes remembering as "as if" perceiving again. This suggests that memory, imagination, and perception share a similar set of property experiences (e.g., Byrne, 2010; Hume, 1739; Matthen, 2010; Nanay, 2016b). Whether you visually see, remember, or imagine an object, you enjoy phenomenally similar experiences of a shared set of colours and shapes. Despite this phenomenal similarity, we normally have little difficulty categorising an experience as a memory, imagination or perception. However, we might be confident about a mental state being in one category but be wrong in our assessment. What accounts for this strong sense of distinctness? While one might say that perception is distinct because it involves no "mental imagery",[1] this cannot be sufficient, since we might mistake mental imagery for perception or vice versa. Our question here is how does the individual distinguish between an experience that is produced by mental imagery, be it memory, imagination or perception, which is not a form of mental imagery? Our answer needs to allow for the fact that we may be wrong about our assessment.

There are three main competing accounts of how memory, perception, and imagination are distinguished. The first, following Hume (1739), appeals to differences in how properties are experienced, such as memories and imagination being less "vivid" than perception. The second, found in James (1892), and more recently in Debus (2016),[2] eschews phenomenology altogether. This view distinguishes memory, perception, and imagination via broader beliefs about our own personal narrative or the production of the state. An experience is recognized as a memory when we acknowledge that it's not produced by the current use of our sensory systems and its content fits within

---

1   See Nanay, B. (2016b) for a counterargument that (much of) perception does involve mental imagery.

2   Perrin et al. (2020, p. 2) cite Redshaw (2014) and Mahr & Csibra (2018) as other defenders of one version of this view. Perrin *et al.* (2020) offer sustained critique of this view.

broader beliefs about our personal past. The third approach says that memories are marked by a feeling of *familiarity* (Russell, 1921) or metacognitive feelings of *knowing* what happened (Dokic, 2014; Perrin, Michaelian & Sant'Anna, 2020). Proponents of this third view often see it as a way of explaining, or cashing out, the feeling of pastness as a metacognitive feeling of familiarity or knowing.

Russell (1921) tentatively proposed that memory is distinguished by a 'feeling of pastness', although this claim may be interpreted in different ways. According to Russell, 'various factors that concur in giving us the feeling of greater or less remoteness in some remembered event' (Russell, 1921, p. 134). One reading of this is that an event can feel more or less temporally distant depending on how long ago it occurred. Our focus here is not the distinction between further and closer temporal events and whether they have a distinct phenomenology[3] but rather the distinction between memory, imagination and perception. As you episodically recall a past-perceived event, it *feels* as if what you are experiencing happened in the past, while imaging and perceiving lack this feeling (see also Matthen, 2010; Perrin *et al.*, 2020; Tulving, 1985). In contrast, what you perceive, but not what you imagine or remember, feels both temporally and spatially *present* (Dokic & Martin, 2017; Matthen, 2005; Windt, 2018). We argue that this picture is not entirely correct; however, the sense of pastness describes an essential aspect of human cognition regarding what it's like to experience memory.

While we argue that the feeling of pastness is a real and important feature of our mental lives, appeals to feelings of pastness and presence have been challenged from two angles. First, Byrne (2010) and Debus (2016) argue that talk of a "feeling of pastness" is mysterious, dubious, or not reflective of anything introspectively found when remembering. Here we respond to these challenges, arguing that the feeling is pastness is a real aspect of remembering. Second, Nanay (2016b) maintains that imagination can involve the feeling of presence. More broadly, supposed feelings of pastness and presence seem to crosscut the divides between memory, perception, and imagination: some perception lacks the feeling of presence, or some memory lacks the feeling of pastness.

We agree that the feelings of pastness and presence neither *define* memory, perception and imagination nor objectively distinguish these experiences. There are too many plausible counterexamples to specify these features as necessary or sufficient conditions.[4] However, we argue they are important aspects of experience that in normal circumstances, are how we categorize each state from a first-person perspective[5] although these feelings can be misleading. The feeling of pastness is what makes a memory —or a false memory, for that matter— feel like a memory. We further

---

3   This is certainly an interesting feature deserving separate analysis.

4   Perrin *et al.* (2020) also discuss the feeling of pastness without presuming that it's necessary or sufficient for memory (assuming only that it's characteristic of memory).

5   Fernández (2020, p. 288) makes a similar claim, but further ties the feeling of pastness to how memories *entitle* us to believe their contents. We do not think the feeling of pastness can do quite that much epistemic, justificatory work.

articulate an explanation of the feeling of pastness as a distinctly phenomenal feeling that is the consequence of the coming together of several cognitive markers. These include spatiotemporal feelings of past embeddedness, objectivity and ergonomic significance (Matthen, 2005), that is, the urge to flinch from objects that might hit you and a feeling that nearby objects can be grasped as well as mental strength, our web of beliefs and probability weightings. This is in contrast to recent work analysing the feeling of pastness as an *epistemic* feeling; of familiarity or knowing (Dokic, 2014; Perrin *et al.*, 2020). We go on to discuss how this feeling of pastness is related to vividness (Hume, 1739) or 'mental strength' (Morales, 2018) and metacognitive beliefs (Debus, 2016; James, 1892).

We begin in section 2 by setting out some of the difficulties determining mental states from first and third-person perspectives, then focus in on relevant philosophical debates in section 3. Then, we analyze the plausibility of the sense of presence and pastness in sections 4 and 5 respectively. After arguing that a cluster of phenomenal and cognitive features likely together contribute to the feeling of pastness in section 6, we evaluate several specific features, including mental strength and belief states. We make the case that 'mental strength', a contested feature, is likely relevant to the sense of pastness. We then argue that a doxastic account of the sense of pastness is not plausible, although belief states likely contribute to the experience.

2 First-personal and third-personal investigations into mental states

When discussing memory, we focus on what is normally called *episodic memory*. When you remember an event, you call to mind a previous experience. This act of episodic remembering strikes you as a reliving or "re-experiencing" of the original experience (Tulving, 2002). During the initial perceptual experience, it introspectively felt as if your experience was presently unfolding before your eyes. You can also close your eyes and imagine an event. These are the sorts of paradigmatic cases we have in mind.

At first glance, it seems that we should simply be able to consult our own experiences to answer what it's like to remember. Historically, introspection has been taken to be reliable, if not infallible (Augustine, 1998; Descartes, 1998; Dretske, 1994). Schwitzgebel (2008) notes, however, that introspection can be unreliable, and we are in fact much better at determining what experiences represent than we are at assessing the experience itself. We can easily, quickly and accurately determine that a tree in front of us has green leaves but trying to determine the nature of this representation —say, what exactly makes it perceptual instead of imagistic— is more difficult. The reliability of introspection seems to depend on the circumstances (Overgaard & Mogensen, 2017) and can be improved with training (Bitbol & Petitmengin, 2013; 2016; 2017), but this alone appears unable to get at what distinguishes paradigm cases of perceiving, remembering, and imagining.

Although we find it easy to categorize mental states when they occur, individuals give different answers regarding what these states are like. People disagree on whether there is a clear distinction in vividness and detail between perception, memory and imagination, for example. This difference in opinion may reflect individual differences in the experience itself. For example, there is broad interpersonal variety in imaginative experience, ranging from complete lack of ability to visually imagine (aphantasia) to highly vivid and realistic imagining (hyperphantasia) (Thomas, 2014; Whiteley, 2020). This may also apply to memory, where people with highly sophisticated autobiographical memory (Palombo, 2018) have much more detailed recall experiences, although it is unknown whether this equates to increased vividness or hallucination-like experience (LePort, Stark, McGaugh & Stark, 2017; Patihis, Frenda, LePort, Petersen, Nichols, Stark, McGaugh & Loftus, 2013). However, considering theorists widely disagree on the phenomenology of these states, neither individual introspection nor empirical analysis has settled the issue.

Aspects of perception can be completed by memory or even imagination. This is important as, although we want to set out what makes it, in most circumstances, easy to categorize a mental state, our confidence about these states does not guarantee accuracy. An example of this is the case of apparent colour perception in the peripheries. Peripheral colour cones are not densely packed enough for vivid or perhaps any colour representation (Mather, 2016), although vivid colour appears to be there, and individuals confidently report perceiving colour. When the eyes saccade to foveate on different parts of the visual scene, this may allow for information about the colour of objects in the peripheries to be stored in memory that is then experienced as perceived colour; a top-down cognitive process that might partially explain the apparent colour (Anstis, 2010). However, we seem to experience colour in patches of the visual field that we did not yet saccade to. Thus a predictive processing explanation in which the mind predicts, or in some sense *imagines* what colour might be there in order to fill in the missing information (Butz, 2017) may also explain part of the experience. Assuming memory, imagination and perception are phenomenally distinct, both explanations involve introspective error. In the first, memory is confused with perception. Further, the illusion does not dissipate when we foveate on one area for an extended period of time. The predictive processing explanation error involves failing to distinguish imagined from perceived colour, and we strongly believe what we are experiencing is perception.

An alternative to introspection is which third-personal methods which analyze mental states. Memory, perception and imagination are deeply intertwined neurobiologically and cognitively (Bone, Ahmad & Buchsbaum, 2020; Horikawa, Tamaki, Miyawaki & Kamitani, 2013; Horikawa & Kamitani, 2017a; 2017b; Penfield & Perot, 1963; Schacter & Addis, 2007). For example, emotional memory overlaps with emotional processing (Arntz, de Groot & Kindt, 2005) and imagining bodily motion, such as playing sports (Mizuguchi, Nakata, Uchida & Kanosue, 2012), has several shared mechanisms with

perceiving. This neural overlap may explain the phenomenal similarities mentioned earlier; however, it cannot help to ascertain that there is no phenomenal difference —we can question whether technology will ever be able to answer questions about phenomenology, but this matter aside, our current neuroimaging technology and methods are certainly insufficiently fine-grained to be able to, say, read perceptual states from brain images. Even if we develop sufficiently fine-grained neuroimaging technology, methods and understanding of neural mechanisms, significant differences between individuals' neuroanatomy and between tokens (Elliott, 2020) may mean that the technology will never be sufficient for determining the phenomenology of these states. Neural and phenomenal overlap does not preclude characteristic phenomenal features by which normal, paradigm cases of imagination, memory and perception are distinguished (Perrin *et al.*, 2020). In typical cases, from a first-person perspective, we often easily make such separations, although when asked to introspect or look to the brain to determine what makes them different, we find no clear answer. The ability to first-personally distinguish mental states is our focus here, even though this ability does not solve the difficult issue of how to define such states; we allow that our first-person categorisations can be wrong. In the following, we outline some of the philosophical debates surrounding this issue.

3 Mental images and experiences

The phenomenology of memory, imagination and perception has brought about much debate amongst philosophers. Here we situate our view in relation to other theorists. Going back to at least Hume, memory and imagination, but perhaps not perception, have been considered experiences of so-called "mental images". We agree with Husserl's point that when we remember and imagine, we don't experience pictures *in the head* but we experience the images as if they are of objects that are *out there* (Jansen, 2010). An imagined tree isn't represented as being inside your head. Just as with perceiving, you experience a tree as somewhere in the distance. You *could* imagine a tree literally growing in your brain, but this is not usually what is meant. Relatedly, it's widely held now by philosophers, and we agree, that imagination and memory share temporal and spatial dimensions with perception (Byrne, 2010; Teroni, 2017). All of these types of experiences have a particular first-person perspective from which you experience items or scenes. Husserl (1975) suggests that we imagine and remember objects not by "conjuring up mental images, which would represent those objects, but by simulating experiences of that object" (Jansen, 2010, p. 144).

Following Husserl, contemporary philosophers widely hold that the *content* of perception, memory, and imagination largely overlaps (e.g., Byrne, 2010; Matthen, 2010). While the content might typically vary in terms of detail or "vividness", more or less the same objects and properties could be experienced from the same spatial

perspective. There need not be anything fundamentally different about what it's like to experience the relevant properties, such as colour and shape regarding vision or timbre and pitch regarding audition, when perceiving, remembering or imagining, although such elements, we argue, on average differ. We think that these differences form part of what leads to our experience as memory, imagination, and perception feeling so distinctive.

In our view, these features come together to form important phenomenal differences: a feeling of pastness, presence or lack of these features. However, these phenomenal features themselves do not form part of the content of the experience. Husserl suggests that what distinguishes perception, memory, and imagination is a certain mode. Imaginative experiences are in a mode of 'inactuality' or 'irreality'. We think this idea of a 'mode' is a plausible way of describing the senses of pastness and presence. For Husserl, the content of imagination, memory and future prospection are all forms of quasi-perception whereas for perception, "the object appears to us, so to speak, 'in person,' as itself present" (Husserl, 2006, p. 18). Imagination lacks this feature, appearing "as though it were there, but only as though" (Husserl, 2006, p. 18). This brings up an important distinction between content and mode, or content and attitude (Matthen, 2010). When you remember last night's sunset, in our view, the feeling of pastness isn't a property, such as 'colour', that your memory attributes to the experience. Rather than the feeling being part of the content, it is part of the way the content is presented.[6] Similarly, Arcangeli (2020) notes that it is important that we don't mistake the content of mental imagery with what she refers to as an "attitude" of mental states where "the attitude is how a mental state represents, the content is what a mental state represents" (p. 307). We think that the sense of pastness and presence are modes or the "how" of presentation rather than part of the content of experience. Our goal here is to not only argue for the plausibility of these modes as key features but also to cash out what accounts for them.

Although it's intuitive to say that imagined objects are experienced as merely *possible* occurrences, here we side with Byrne who thinks this is wrong. According to Byrne, "what 'appears to be so', when one imagines a purple polar bear, is that purple polar bears exist, not (merely) that they could have existed. (Of course, one will not believe that things are as they appear)" (Byrne, 2007, p. 135). While we agree that the occurrence of the imagination of the bear appears to be so now, an extra layer is added for objects of memory, the sense of pastness, and perception, sense of presence. You might suggest that when we remember past events, just as with imagination, we experience them *as happening* —we experience, or at least *re-experience* (Tulving, 2002) them as happening now— not as *having happened.* The replaying of the memory of last night's

---

6   Dokic (2014) and Perrin *et al.* (2020) outline views where the feeling of pastness is a metacognitive element that is not part of the content, nor the attitude taken toward that content, but results from monitoring the memory. In our view, there are a series of components beyond these features, such as mental strength, that are plausible components that are causal factors that lead to a feeling of presence.

sunset or imagining a purple polar bear is experienced as occurring in the *present.* In our view, while the contents of the mental state of all three, perception, memory and imagination, do feel as if they are happening now, only the events represented by the mental states of perception appear as if they are currently occurring, only they have the sense of presence. These features are lacking in imagination. How do we then distinguish between, say, imagining something currently occurring or imagining something happening in the past? Imagination lacks these senses of pastness or presence, as we argue, otherwise, it could be mistaken for perception or memory. The distinction between imagining something currently occurring or having occurred might be distinguished by, say, doxastic features; beliefs about what the imagined scenario would be like or would have been like. We discuss why doxastic features are not plausible for the key distinction between memory, imagination and perception in section 8. However, first, we will formulate our account of the relevant feelings.

In order to argue that the senses of presence and pastness are important features of our mental life, we will discuss each in turn. We turn to an account of presentness first, which is, perhaps, more widely accepted than the feeling of pastness. The key features of presentness will help explain some of the similar features of pastness.

4 Feeling of presence

Perception is often said to be accompanied by a 'feeling of presence', that objects and events are experienced as occurring *here and now* (Dokic & Martin, 2017; Matthen, 2005; Windt, 2018). As Dokic & Martin (2017) point out, it's been noted at least since Husserl that perceptual experience strikes us, introspectively, as if we are experiencing the present environment around us. We argue that this sense, and the sense of pastness, are not in fact too mysterious to be useful (Byrne, 2010; Debus, 2016): they are helpful for distinguishing mental states from the first-person perspective.

Let's say you are walking through the forest, perceiving the trees surrounding you. This feeling has several dissociable components. First, you feel immersed, that you are *in* a forest rather than looking *at* it (Revonsuo, 1995; Windt, 2010) as you would if you were looking at a picture of a forest or imagining being in a forest. Second, objects feel ergonomically significant: You can reach out and grab a tree branch and feel the need to duck out of its way as you pass. Third, there is the feeling of objectivity or reality, that perceived objects, trees, rocks and so forth, unlike, say, phosphenes, are mind-independent (Siegel, 2006). Fourth, perceptual experience of objects strikes us as 'accessible': our own body movements as we walk around a tree can bring new, previously unseen, parts into view (Noë, 2004).

While everyday perceptions often involve all four components of the feeling of presence, unsuccessful perceptual experiences or hallucinations may lack some of them while maintaining others. We might assume that dreams, immersive hallucinations,

or sensations brought about by artificial sensory stimulation would involve the full feeling of presence, but this is not always the case. A dreamed forest may seem under-represented, vague and unstable. Memory-based hallucinations induced by direct temporal-lobe stimulation during clinical seizure interventions seem to lack important components. Penfield and Perot (1963) report that patients who had audiovisual hallucinations during surgery never lost track of the actual operating room, nor mistook their hallucinations for real perception or felt an urge to interact with them, suggesting a lack of ergonomic significance. The common description of these hallucinations as "dream-like" suggests they lacked the felt sense of objectivity. Penfield's patients often reported auditory hallucinations that sounded as if they were heard through a telephone or radio, suggesting inaccessibility.

Still, these temporal-lobe hallucinations seem to at times involve the feeling of immersion. One of Penfield's patients reported that "I see the people in this world and in that world too at the same time" (Penfield & Perot, 1963, p. 679), suggesting that they felt immersed in both their genuine perceptual and hallucinatory experiences. Revonsuo (1995), Windt (2010) and Metzinger (2009) have argued that dreams also involve the feeling of immersion, although it's plausible that many dreams lack this and other components of presence (Barkasi, 2020a; 2021; Rosen, 2013; 2018b). Dreams that occur in non-rapid eye movement sleep (NREM) are likely to be more imagination-like than hallucination-like (Rosen, 2018a; 2019) with rapid eye movement sleep (REM) dreams usually described as more immersive and multisensory (Schredl, 2018). However, there is disagreement about whether any dreams should be described as perceptual (Ichikawa, 2008; 2016; Ichikawa & Sosa, 2009) or if so, whether they are convincing, real world-like hallucinations (Noë, 2007). Typical imagination, closing your eyes and visualizing a red apple, lacks *all* of these components (Dokic & Martin, 2017; Matthen, 2005). It is plausible that many drug-induced, Lilliputian, and visual-release hallucinations also lack them, given that they are often described as seeing a strange object floating in the visual field, but not seen as out in the world (see the descriptions in Ffytche, 2013). If waking hallucinations and dreaming were a form of imagination (e.g., Allen, 2015; Ichikawa, 2008; 2016; Ichikawa & Sosa, 2009; Nanay, 2016a), we would expect a strong overlap with typical imagination, lacking a sense of presence.

Perception may not involve all aspects of the feeling of presence, even in successful, veridical cases. For example, mountains seen in the far distance are not experienced as ergonomically significant or accessible via small bodily movements. One could argue that the experience *as a whole* is still immersive since the mountains are only a far-away part of the visual scene that you are immersed in, but elements of the visual scene can certainly feel more or less present. Some types of virtual reality (VR) may be good examples of perceptual experiences lacking the feeling of immersion and reality. Generating these feelings is a major aim of VR (Grassini & Laumann, 2020; Sanchez-Vives & Slater, 2005), but bad setups fail in this goal. Thus, causally deviant situations

in which the normal perceptual flow of information is disrupted can break the feelings of immersion and reality in an otherwise successful or genuine perceptual experience.

Experiences of real sensory stimuli can therefore lack the feeling of presence, while aspects of the feeling of presence can be involved in imagination. So, the feeling of presence isn't fundamental to perceptual experience. It is more plausible to say that the feeling of presence is what makes perception feel like perception *rather than* what in fact makes an experience perceptual, and, further, causes us to mistake other mental states for perception. A perceptual experience that somehow loses its sense of presence remains perceptual but may register, from the first-person perspective, as imagination.

The sense of pastness, like the sense of presence, is not a part of the content of memory but rather a mode under which this content is experienced. In contrast, there are a separate set of cognitive features that distinguish a sense of pastness from a sense of presence as we discuss in the following section.

5 Feeling of pastness

Understanding the elements of the sense of presence from the previous section will help to understand the sense of pastness. Tulving's (1983) talk of "autonoetic consciousness", the view that there is a "special kind of consciousness that allows us to be aware of subjective time in which events happened" (Tulving, 2002, p. 2) evokes a similar idea to the feeling of pastness. Some theorists deny that the sense of pastness is a phenomenal feature of memory (Byrne, 2010; Debus, 2016), although we take a different view on this; sense of pastness is what makes a mental state feel like memory. We have two goals here, firstly to argue that this feeling is what makes it so easy for us, from a first-person perspective, to class a mental event as a memory. This contrasts with views that deny the relevance of a feeling of pastness all together (e.g., Byrne, 2010, Debus, 2016) and those that consider feeling of pastness as a defining feature of memory, say, under a non-factive view of memory (Fernández, 2019). This is because we want to allow that individuals can mistake other experiences for memory, which, for us, is a more plausible explanation of how mental states overlap and interrelate. A misplaced feeling of pastness is, in our view what explains this mistaken identity. The feeling of pastness works as a useful heuristic, although it is not always accurate. This also explains the common occurrence of knowing a mental event is not a memory but feeling that it is. Realising that an event did not occur does not dispel the feeling of memory precisely because such beliefs do not dispel the feeling of pastness. Our second goal is to evaluate possible cognitive features that underly the sense of pastness. This is discussed in detail in the following sections.

As the sense of presence, the sense of pastness, when considered as a mode of presentation, is not a mysterious feature. When remembering walking through the forest, the scene has a sense that it *was* immersive. It feels like if you *had* moved

in a certain direction, the backs of objects would have been accessed. The objects *were* ergonomically significant but no longer are. Might a memory instead involve a current sense of ergonomic significance? When remembering turning a light on and thinking about turning it off again, the light remains there to be manipulated, and perhaps *feels so*.[7] We think that there is a sense in which ergonomic significance is maintained but altered. Now the switch feels present but, say, in another room, not in front of you as represented in memory. The light switch feels *reachable* only by walking back to the other room, not reachable in your *peripersonal space.* If it was represented as reachable, the experience might be misinterpreted as perception, a hallucination. There might be some overlap if you switched on the light then closed your eyes and remembered the light switch. The switch does feel present because you could reach out and touch it, but the memory of it has a sense of pastness. There are some complexities, but these senses aren't truly mysterious. Another important issue is that these senses do not necessarily accurately determine the mental state, they can lead to miscategorising.

We can mistake memory for imagination when it lacks a sense of pastness, or, perhaps less commonly, mistake memory for perception (Barkasi, 2020b) or imagination for perception. For example, imagining a grating at a specific orientation makes subjects more likely to report detecting that grating (Dijkstra, Mazor, Kok & Fleming, 2021). This, in our view, is explained by the image gaining a sense of presence due to relevant cognitive features going awry. What makes imagination feel different from perception and memory is that it lacks a sense of presence or pastness but since there are overlapping cognitive attributes, they can be confused. In our view, a memory that loses its feeling of pastness remains a memory but is experienced as imagination whereas if a sense of presence is gained, it would feel like perception but not be ontologically classed as such. Further, the phenomenal features of any particular memory, perception or imagination may differ depending on the individual and circumstance. It is likely that some or all of these features lie on a spectrum so that, like vividness, past ergonomic significance can be more or less intense, one might have a stronger or weaker sense of *having been* immersed, and so on. These mental features may not 'carve nature at its joints', yet the clear sense of pastness and presence or absence thereof may often make it very easy for an individual to feel that "this experience is a memory".

If sense of pastness is what makes a mental state feel as it does but does not define memory, this is consistent with empirical findings that suggest that presentness and pastness cross-cut the lines between perception, memory, and imagination. Nanay (2016b) denies that the feeling of presence distinguishes perception and imagination since certain imaginative experiences can involve the

---

7   Thanks to a reviewer for this example.

feeling of presence. Similarly, Debus argues that "one might plausibly suggest that there are at least some cases of R-memory[8] which are not accompanied by relevant feelings [of pastness], and one might well come to hold that subjects do not usually experience any relevant 'feelings of pastness' when they R-remember" (Debus, 2016, p 138). We agree that the feelings of presence and pastness can come apart from perception and memory. But from this, we should neither infer that there are no feelings of pastness or presence, nor that these feelings fail to serve as markers for the mental state types.

The feeling of pastness is not what makes an experience a memory but only what makes it feel as such. An experience of imagination may still involve some, but not all or at sufficient intensity, components of the feeling of presence or pastness, but by passing a threshold can be mistaken for another mental state. Any state may be falsely classed as a different type. These feelings, however, are generally useful heuristics that lead to easy categorisation but can be cognitively superseded, e.g., by evidence that their 'memory' is false.

In contrast, it might be argued that imagination that gains a sense of presence should then be *classed* as hallucinatory perception, or that experiences lacking the feeling of pastness are no longer memories. We find this implausible, as it requires radically redescribing or re-taxonomizing all the cases discussed above. We think a broader criterion of mental state types which considers phenomenology, neurobiology, functional role, and causal-contextual-etiological factors should be adopted. How these come together is an important issue for further research. If an experience is produced from a memory trace with causal origins in a past event but lacks a sense of pastness, we would still call this a memory.

Our view is compatible with contemporary constructivist and anti-causal views of memory (De Brigard, 2014; Michaelian, 2011; 2016; Robins, 2019; Sant'Anna & Michaelian, 2019; Sutton, 1998). It's not etiology alone which defines memory, but some mix of etiology, neurobiology, functional role, and phenomenology. We assume that even constructivists will agree that there's more to what makes a state a memory than just its phenomenology. We do, however, want to allow for a classification of 'false memories' that are *not* memories but, instead, are other types of cognition mistaken for memory. This contrasts with Fernández's (2019) non-factive approach, according to which false memory is a type of memory because the term 'false memory' does not seem to be "a deviant use of the term "memory." Which it should, if "memory" was factive" (Fernández, 2019, p. 3). Under a non-factive view, memory ontology *could* be determined by phenomenology; mental states that feel like memories, *are.* The sense of pastness could be the defining feature. We, however, think it is more

---

8   'R-memory' is Debus' term for 'recollective memory', or a memory with an experiential component, as contrasted with mere propositional recall.

plausible that false memories can be miscategorised imaginations, as we discuss further in the following.[9]

Perhaps here there is only disagreement in approach, as Craver's (2020) suggests. The empirical approach takes memory to be fallible while the epistemic approach takes memory to be a way of accurately representing the past and gaining knowledge about it. If, however, we can be wrong about introspection, even under an empirical approach, it is still possible to be wrong about memory under a memory-as-phenomenology view. We might, in a difficult case, introspect and make a wrong assessment about the phenomenology of the mental state, assessing that 'this feels like it happened in the past' when it does not. If phenomenology picks out the type of mental state but we can be wrong about phenomenology, we can be wrong about the mental state type.

We find it more plausible to deny that false memory is a type of memory. The sense of pastness generally does a good job at helping us to be in touch with the past; picking out mental events that involve mental time travel (Barkasi & Rosen, 2020). We can, however, mistake imagination for memory when the sense of pastness is associated with the wrong type of mental event. On the other end of the spectrum, however, we might even be wrong about a feeling of pastness, a mistake made through faulty introspection, leading us to be wrong about something even feeling like a memory.

One issue is that perception, memory, and imagination overlap and intermingle in a way that conflicts with distinct phenomenology. In Levy's (2012) interpretation of Sartre, memory "supplies materials" to imagination and "imagination shapes our memories" (p. 156), which is plausible given that imagining a face is informed by our memory of what that face looks like and that memory is always, to a certain extent, reconstructed using imagination (Sutton, 1998). While plausible in terms of the intertwining of cognitive features, Levy (2012) goes a step further to argue that "imagination and memory cannot be separate from each other" (p. 156). We think that, although imagination and memory can be mistaken for each other, there are clear phenomenal features that make them usually easily distinguishable from a first-person perspective yet which cannot be used to categorize them ontologically.

Now that we have set out sense of presence and pastness as modes of experiencing mental content, we argue in more detail why several of the cognitive features discussed by other theorists should be seen as part of a cluster of indicators that comprise these senses. These lead to an experience having a mode of presentness, pastness or lacking either feature.

---

9   False memories are presumably cases of imagination confused for memory. If hallucination is a form of imagination (e.g., Allen, 2015; Nanay, 2016a), then it's a case of imagination confused for perception. For example, failure of self-monitoring can lead to mistaking internal monologue for voices in schizophrenia (Bob & Mashour, 2011). The famous Perky experiment involved perception being confused for imagination (Perky, 1910).

6 Causal factors of pastness and presence

Previously we responded to those who argue that pastness and presence are too mysterious to be useful (Byrne, 2010; Debus, 2016) by analysing these concepts. Now we discuss what causal factors might lead to these senses.

6.1 Basic phenomenal and cognitive features

We think that several more basic phenomenal and cognitive features contribute to the sense of pastness. For example, a feeling of familiarity (Dokic, 2014; Russell, 1921) may be part of the experience. 'Mental strength', a much-contested aspect of experience, may also contribute while not being the key feature that distinguishes mental states. To explain, a mental state might have a stronger sense of pastness, thus feel more like a memory, if it is more vivid than the average imagination and less than the average perception. We may associate recalled events with being less vivid due to the passage of time as objects that are spatially further away are also less vivid.[10] Should older memories, therefore, have less of a sense of pastness? There may not be symmetry between presence and pastness in this regard, as it seems that although objects that are spatially further away feel less present, objects temporally further may not feel 'less past' but just 'further past'. However, it is likely that often, older memories do become ambiguous and lose vividness; it then becomes harder to discern whether a memory from long ago is a memory or simply imagined. Further, a background web of beliefs may contribute to the feelings of presence and pastness. We discuss mental strength in more detail in section 6.2 and beliefs in 6.3.

A good analogy for this 'multiple indicators' model is how a cluster of 'clues' contribute to a sense that a perceived sound emanated from a particular direction. These include time lag, distinct volume and quality of sound between ears as well as cues from other senses such as vision (Mather, 2016). From a first-person perspective, we cannot specify *why* a sound seems to emanate from a direction —some features are not consciously detectable, such as the imperceptible difference in volume between ears, and others go unnoticed. Yet together, these features bring about the phenomenal feel of the sound coming from *over there.* Similarly, we might not be able to consciously distinguish between the vividness of memory and imagination, but they may contribute to the sense that an experience comes from the past from below the level of awareness.

---

10   Memories may also have a sense of *having been* more or less vivid. If remembering a hazy or unvivid perception, that original experience will likely carry forward to the experience in the sense that the memory is of a perception that was hazy. An interesting attribute here is that in memory, detail could be added via imaginative filling in, making the memory clearer than the perception, but we discuss this issue in the following section.

Experience results from noisy and ambiguous input (Pizlo, 2001). Perception, memory, and imagination fill in for each other and intermingle (Pelaprat & Cole, 2011) such that our general experience relies on each capacity functioning well (Mitterer, Horschig, Müsseler, & Majid, 2009). This can cause errors, such as occurs with hypervigilance; misinterpreting neutral, ambiguous stimulus as being threatening due to memory or preconceptions (Kimble, Boxwala, Bean, Maletsky, Halper, Spollen & Fleming, 2014).

If filled-in detail and perceptual detail are indistinguishable, perhaps this simply is what perception *is* —combined sensory feedback and top-down modulation intermingling with gap-filling that draws on memory and imagination (Albright, 2012; Penfield & Perot, 1963). If experience is just intermingling of sources and cognitive features, then the *ontological* boundary between memory, perception and imagination becomes less clear. But this ontological messiness is consistent with overall distinguishable feelings of presence and pastness. These feelings often seem to supervene on diverse mixtures of neurobiological machinery, cognitive and phenomenal features. Despite the cognitive intermingling, there does seem to be a clear sense of presence assigned to certain mental states and pastness to others. It is plausible that this is an outcome of the weighing up of these features, many of which work as cues that are below the level of conscious awareness.

We conclude that the feelings of pastness and presence are phenomenal markers of, respectively, memory and perception, while imagination lacks these features. This is consistent both with the fact that some tokens of memory and perception, typed by a more holistic criterion including etiology, neurobiology, and functional role, lack these features. It's also consistent with the intermingling of perception, memory, and imagination state types. In the following, we focus more specifically on the relevance of two different potential cognitive contributors, mental strength and beliefs.

6.2 The contribution of mental strength

Historically, vividness, clarity, detail, intensity, strikingness and other features described under the umbrella of 'mental strength' by Morales (2018) have been taken as markers of mental state type. Hume famously proposed a difference in the degree of "vivacity". While memory "retains a considerable degree of its first vivacity, and is somewhat intermediate betwixt an impression and an idea", imagination "entirely loses that vivacity, and is a perfect idea". Perception is more vivid than imagination and memory, although he acknowledged that this didn't hold true for each occurrence of a mental state (Hume, 1739, II:5).

These features are generally rejected by modern researchers as *fundamental* distinctions (Debus, 2016; Teroni, 2017), while others reject their coherence entirely (Kind, 2017). Perceptual experiences of properties can be quite "dim", while memory

and imagination experiences can be hyper "vivid" (Perky, 1910; Thomas, 2014). We deny this wholesale rejection, however. Mental strength likely relates to the feelings of pastness or presence. While some perceptions can be dull and some imaginations vivid, generally imagination is less vivid than perception. The presence of an object allows visual and other sensory exploration and details are made available to attention. In contrast, imagination is more likely to be under-represented, that is, an imagined scene may lack detail or objects. The mental effort required to vividly represent, say, imagining walking down the street with the level of detail and complex interaction of multiple senses that occurs in perception, requires significant cognitive resources. Since the experience is generated by the mind, whereas perception involves taking in sense data from the surrounding environment, it is difficult, if not impossible, to imagine a scene with as much detail as a perceived scene. Memory may be more vivid on average than imagination since representations are to an extent retrieved from memory traces rather than generated. Just noticeable difference, a major foundation of cognitive science of perception (Mather, 2016), shows that we can distinguish between different subtle features of weight, illumination, saturation and so forth. We can also judge if something appears more or less blurry, saturated and detailed. Judgments about features of perceived objects tend to be more reliable than introspective judgments. There certainly seems to be *something* to talk of mental strength. Mental strength may, *in part*, furnish what makes mental states, as previously discussed, so easily distinguishable from the first-person perspective in most cases.[11] This is supported by the fact that mental strength can make real objects feel more or less present.

A strong image can be more intense than a relatively weak perception. Vision in low light, mist, in the distance, for someone with poor eyesight or any of the many circumstances where vision is impaired are examples. A perceived object might be indistinguishable from an image in extreme circumstances, although we might still be able to distinguish perception from imagining or remembering despite not knowing what it is we are looking at. Relevant cognitive features likely contribute to a threshold of sense of presence where one will confidently state that an experience is perceptual. Less presence leads to less confidence. Someone who enters a room without their glasses on might find that the blurry objects feel *less present* than when they later don their glasses. A blurry colour patch can be indistinguishable from an after-image. One could argue that, in contrast, when a person takes off their glasses and their vision becomes blurry, objects to them *don't* lose their sense of presence, they retain the sense that they are there but just look blurry.

---

11 Interestingly, a virtual reality environment of textureless line drawings can elicit the same physiological responses as one with more detailed, realistic renderings (Sanchez-Vives & Slater, 2005). This shows that some mental strength markers, like detail, don't correlate well with the feeling of presence (which itself seems to be part of how we normally distinguish perception from memory and imagination).

This could be a case of memory or imagination making up for the lack of detail, allowing for sense of presence to remain stable. When observing the face of a loved one very briefly, details such as the small mole on their cheek may be filled in as we fill in the colour in our peripheries. Remembering the detail of a well-known object might assist in maintaining its sense of presence. Blurry objects may also have less of a sense of availability and accessibility since it is unclear how the object can be interacted with, but changing one's perspective may disambiguate the object, increasing accessibility and ergonomic significance. Similarly, when one is trying to remember what happened at lunch yesterday, imagination may help fill in the blanks while perception, let's say, being in the same room that one was in during that lunch, may fill in some of the details that weren't originally remembered. It is likely that particular aspects accounting for the experience are often unknown by the individual undergoing the experience.

What is required to pass a 'presence' or 'pastness' threshold likely differs between individuals. Individuals with poor eyesight may adapt to rely less on mental strength, depending more heavily on other features to attain a sense of presence. Further, someone who closes their eyes might feel that an unseen object is still present. It can be argued that presence is related to mind independence, however, if one has their eyes closed and is told there is an object in front of them, this may not have the same sense of presence even if the belief is formed that there is indeed an object there.

So, it is likely that belief in a mind-independent object can add to a sense of presence but is not sufficient. If it is possible for a realistic hallucination to involve a sense of presence, as may be so for lucid dreams, where the dreamer realises they are dreaming (LaBerge, 1981), then belief of mind-independence is not necessary either. Mental strength is likely a part of a cluster of causal features that make an experience feel the way it does and thus should not be overlooked.

A final feature we analyse as part of the indicators that lead to experiences of pastness and presence is belief states. In the following, we argue against a doxastic view of the distinction between imagination, perception and memory, but note that beliefs play an important role in our experience.

6.3 The relevance of belief to phenomenology

An alternative view to ours that comes from James (1892) and other contemporary theorists (e.g., Debus, 2016), is that memory is distinguished from imagination via biographical knowledge, a broader set of beliefs rather than phenomenology. Your current experience might match your beliefs about an event from the past, and not your belief that, say, you're now sitting at home. On that basis, you infer that the experience is a memory. In contrast to doxastic views of experience, in our view, the sense that something is a memory does not depend on beliefs about the actual presence of the experienced scene, but beliefs may be relevant to the sense of pastness. Just as I

continue to see an illusion despite knowing that it is an illusion, a hallucination could have a sense of presence despite the viewer knowing that it is unreal.

Belief is likely relevant to the experience of memory and may modulate phenomenology. While holding a specific belief does not usually alter phenomenology, the right *network* of tacit beliefs might contribute towards probability weighting that induces the feeling of pastness described above. Higher-order processes modulate and alter lower-level ones. For example, probability weighting can affect whether a concave mask is seen as convex (Corlett, Taylor, Wang, Fletcher & Krystal, 2010). While one's current belief that the mask is concave does not cause it to appear concave, a tacit web of beliefs that faces are convex may contribute to the illusion. Probabilities relating to the context of an image can also affect the perception of its colour, as occurs with 'the dress' illusion, which some see as blue and black and others as white and gold (Handel, 2019) or with sound, as with the McGurk effect (McGurk & MacDonald, 1976), where seeing a mouth move a certain way alters how an ambiguous word sounds. This does not mean, however, that the content is not itself phenomenal. Belief or probability weighting may intensify the experience by heightening the related phenomenal attributes such as current or past embeddedness, accessibility and mental strength. "Presentness" and "pastness" describe an experiential mode rather than doxastic formation although they may be modulated by one's web of beliefs or probability weightings. For example, the top-down modulation that leads to 'the dress' being either black and blue or white and gold occurs without explicit awareness of a belief about background luminosity. Individuals are surprised to find that others see the dress differently and the explanation about interpreted context is not obvious. Beliefs and phenomenology about our experience of an object being *present* or *in the past* seem hard to disentangle. While beliefs can influence phenomenology, it is plausible that the feeling of presentness or pastness can lead to the formation of belief, say, about the veracity of the experience and when it occurred.

Importantly, beliefs and feeling here are dissociable. We can phenomenally feel as if what we're experiencing is present or past without believing it is, and vice versa. For example, while lucid dreaming, realising that one is dreaming doesn't necessarily make the phenomenology less realistic (Metzinger, 2003; Revonsuo, 1995). In fact, some describe lucid dreams as being more vivid and realistic than non-lucid dreams (LaBerge, 1985; 2000), although this may be due to other cognitive attributes such as increased attention and memory (Filevich, Dresler, Brick & Kühn, 2015; Voss, Schermelleh-Engel, Windt, Frenzel & Hobson, 2013). That is not to say that lucidity cannot or never dispels the sense of presence, just that there isn't a necessary or tight connection between the two. This view leaves open the disjunctive interpretation that hallucinations, although convincing, are only perceptual from the first-person perspective but are not in fact perception. This would be analogous to considering false memories as non-memories.

Similarly, a VR experience can and often does successfully induce a feeling of presence despite the subject knowing that it is merely virtual. Just as you continue to see illusions, such as the Müller-Lyer, despite knowing that it is an illusion, it is possible for a hallucination to have a sense of presence despite the viewer knowing that it is unreal. How people interpret the background affects the perception of the colours of the stripes of 'the dress' illusion in a top-down fashion, but the process goes beneath conscious awareness just as with sensing where a sound came from. Perhaps belief at some tacit level or instead, a predictive weighting as described by predictive processing, affects perception (Weise & Metzinger, 2017). We argue, however, that the experience of pastness is phenomenal and we categorize based on phenomenology rather than belief. Feelings of presence or pastness, however, may *relate* to beliefs about the actual presence or pastness of the experienced scene. While presence and pastness are genuine phenomenological features, beliefs and phenomenology are deeply related.

Debus (2016) mentions the fickleness of feelings and how a feeling of pastness should not be sufficient for an individual to form a belief that their memory is accurate. This does not strike us as a decisive objection. Reflections on accuracy may be a process that occurs *after* the initial experience of the mental event; the experiencer feels that their mental state is either a memory or perception and they may then go on to reflect about whether their feeling is accurate. We think the question of what makes an experience feel like memory is different from the question of how we assess the accuracy of a memory.

Although the doxastic distinction between memory, imagination and perception allows for there to be false memories and to mistake one type of experience for another, it is more plausible that the phenomenal senses of presence and pastness rather than doxastic features are what make the experience of perception and memory what they are. One argument for this is that a mental state can feel like memory without the individual believing that it is. For example, one assesses that a mental state that feels like a memory cannot possibly have happened, based on external evidence. The belief doesn't change the feeling. Secondly, mental states are often categorized more quickly than it would take to assess a belief about the state, for example, when you duck from an incoming projectile, the sense of presence may proceed any explicit belief formation. We may accept a memory as such based on phenomenology without rational assessment, or form a belief about a memory based on the feeling of pastness alone.

The doxastic view seems to provide a simple explanation for why false memories can be relatively easily brought about in experiments (e.g., Loftus & Pickrell, 1995; Wade, Garry, Read & Lindsay, 2002). If there is no particular memory-phenomenology to distinguish the states, all that would be required is to believe one's image was memory for it to be classed as a memory. However, this is not a reason to reject the importance of the sense of pastness in our mental lives. Firstly, a sense of presence may develop over time in the experiments mentioned above. Secondly, we also have beliefs about memories being more or less accurate of the original event that contrast with how the

mental events *feel.* In normal circumstances, at first pass, we don't assess or judge an experience as falling into one category or another; we experience it as such. This accounts for the ease and confidence with which we initially distinguish between states, irrespective of accuracy. Sense of presence and pastness are their own phenomenal features despite having a complex relationship with belief. Belief is relevant to these processes, as described. But it is primarily a sense of pastness that makes us ascribe an experience as memory rather than what defines memory, allowing for mistaken identity.

Conclusion

The feeling of pastness is a good contender for what makes a mental state *feel* like memory and part of what gives us confidence in memory. A strong sense of pastness is likely what makes it easy, in most circumstances, to quickly judge "I remember" as distinct from "I imagine" and "I perceive" without reflection. At the same time, sense of pastness can be inaccurate, and we can be wrong about mental states that strongly feel like memories.

We argue that the feeling of pastness is a truly phenomenal aspect of the human experience and that this feeling is strongly related to a complex interaction between cognitive components including a sense of space and time. We agree with theorists who deny that the sense of pastness and presence define and distinguish imagination, memory and perception. However, these senses are plausible contenders for what makes a particular experience feel the way it does. Specifically, the *feeling of pastness* consists in a felt relation between our current structured specious present and an experienced scene, of which we don't seem to be a part, that lies in relation to our current specious present. We argued that the feeling of pastness is influenced by both phenomenal features such as 'mental strength' and feelings of *having been* immersed and objects *having had* ergonomic significance, and metacognitive doxastic states such as beliefs about a mental state's fit with the biographical narrative. Feelings of pastness and presence are fallible heuristic markers by which we identify experiences as perceptions, memories, or imaginings. Despite not solving the ontological issue of mental states, these are important features of cognition that allow us to pre-reflectively experience and first-personally classify these mental states the way we do.

References

Albright, T. D. (2012). On the perception of probable things: neural substrates of associative memory, imagery, and perception. *Neuron, 74* (2), 227–245. https://doi.org/10.1016/j.neuron.2012.04.001

Allen, K. (2015). Hallucination and imagination. *Australasian Journal of Philosophy, 93*(2), 287–302. https://doi.org/10.1080/00048402.2014.984312

Anstis, S. (2010). Visual filling-in. *Current Biology, 20*(16), R664-R666. https://doi.org/10.1016/j.cub.2010.06.029

Arcangeli, M. (2020). The two faces of mental imagery. *Philosophy and Phenomenological Research, 101*(2), 304-322. https://doi.org/10.1111/phpr.12589

Arntz, A., de Groot, C. & Kindt, M. (2005). Emotional memory is perceptual. *Journal of Behavior Therapy and Experimental Psychiatry, 36*(1), 19–34. https://doi.org/10.1016/j.jbtep.2004.11.003

Augustine. (1998). *The city of god against the pagans.* Cambridge University Press. https://doi.org/10.1017/CBO9780511802300

Barkasi, M. (2020a). Does what we dream feel present? Two varieties of presence and implications for measuring presence in VR. *Synthese*, 1-27 [Preprint]. https://doi.org/10.1007/s11229-020-02898-4

Barkasi, M. (2020b). Some hallucinations are experiences of the past. *Pacific Philosophical Quarterly, 101*(3), 454-488. https://doi.org/10.1111/papq.12320

Barkasi, M. (2021). What should the sensorimotor enactivist say about dreams? *Philosophical Explorations, 24*(2), 243-26. https://doi.org/10.1080/13869795.2021.1908575

Barkasi, M. & Rosen, M. G. (2020). Is mental time travel *real* time travel? *Philosophy and the Mind Sciences, 1*(1), 1-27. https://doi.org/10.33735/phimisci.2020.1.28

Bitbol, M. & Petitmengin, C. (2013). A defense of introspection from within. *Constructivist Foundations, 8*(3), 269–279.

Bitbol, M. & Petitmengin, C. (2016). On the possibility and reality of introspection. *Mind and Matter, 14*(1), 51-75.

Bitbol, M. & Petitmengin, C. (2017). Neurophenomenology and the microphenomenological interview. In S. Schneider & M. Velmans (Eds.), *The blackwell companion to consciousness,* 2 edition (pp. 726-739). Wiley & Sons. https://doi.org/10.1002/9781119132363.ch51

Bob, P. & Mashour, G. A. (2011). Schizophrenia, dissociation, and consciousness. *Consciousness and Cognition, 20*(4), 1042-1049. https://doi.org/10.1016/j.concog.2011.04.013

Bone, M. B., Ahmad, F. & Buchsbaum, B. R. (2020). Feature-specific neural reactivation during episodic memory. *Nature Communications, 11*(1945), 1-13. https://doi.org/10.1038/s41467-020-15763-2

Butz, M. V. (2017). Which structures are out there: learning predictive compositional concepts based on social sensorimotor explorations. In *PPP-Philosophy and Predictive Processing* (pp. 1-16). MIND Group.

Byrne, A. (2007). Possibility and imagination. *Philosophical Perspectives, 21*(1), 125-144. https://doi.org/10.1111/j.1520-8583.2007.00123.x

Byrne, A. (2010). Recollection, perception, imagination. *Philosophical Studies, 148,* 15-26. https://doi.org/10.1007/s11098-010-9508-1

Corlett, P., Taylor, J. R., Wang, X.-J., Fletcher, P. C. & Krystal, J. H. (2010). Toward a neurobiology of delusions. *Progress in Neurobiology, 92*(3), 345-369. https://doi.org/10.1016/j.pneurobio.2010.06.007

Craver, C. F. (2020). Remembering: epistemic and empirical. *Review of Philosophy and Psychology, 11*(2), 261-181. https://doi.org/10.1007/s13164-020-00469-7

De Brigard, F. (2014). Is memory for remembering? Recollection as a form of episodic hypothetical thinking. *Synthese 191,* 155-185. https://doi.org/10.1007/s11229-013-0247-7

Debus, D. (2016). Imagination and memory. In *Routledge Handbook of the Philosophy of Imagination* (pp. 135–148). Routledge.

Descartes, R. (1998). *Discourse on method and meditations on first philosophy* (D. A. Cress, Trans.). Hackett Publishing Company.

Dijkstra, N., Mazor, M., Kok, P. & Fleming, S. (2021). Mistaking imagination for reality: congruent mental imagery leads to more liberal perceptual detection. *Cognition, 212,* 104719. https://doi.org/10.1016/j.cognition.2021.104719

Dokic, J. (2014). Feeling the past: a two-tiered account of episodic memory. *Review of Philosophy and Psychology, 5,* 413-426. https://doi.org/10.1007/s13164-014-0183-6

Dokic, J. & Martin, J.-R. (2017). Felt reality and the opacity of perception. *Topoi, 36*(2), 299-309. https://doi.org/10.1007/s11245-015-9327-2

Dretske, F. (1994). Introspection. *Proceedings of the Aristotelian Society, 94,* 263-278. https://doi.org/10.1093/aristotelian/94.1.263

Elliott, M. L., Knodt, A. R., Ireland, D., Morris, M. L., Poulton, R., Ramrakha, S. & Hariri, A. (2020). What is the test-retest reliability of common task-fMRI measures? New empirical evidence and a meta-analysis. *bioRxiv* [preprint] https://doi.org/10.1101/681700

Fernández, J. (2019). *Memory: a self-referential account.* Oxford University Press. https://doi.org/10.1093/oso/9780190073008.001.0001

Fernández, J. (2020). Self-referential memory and mental time travel. *Review of Philosophy and Psychology, 11,* 283–300. https://doi.org/10.1007/s13164-019-00453-w

Ffytche, D. H. (2013). The hallucinating brain: neurobiological insights into the nature of hallucinations. In F. Macpherson & D. Platchias (Eds.), *Hallucination: Philosophy and Psychology* (pp. 45–64). The MIT Press. https://doi.org/10.7551/mitpress/9780262019200.003.0003

Filevich, E., Dresler, M., Brick, T. R. & Kühn, S. (2015). Metacognitive mechanisms underlying lucid dreaming. *The Journal of Neuroscience, 35*(3), 1082-1088. https://doi.org/10.1523/JNEUROSCI.3342-14.2015

Grassini, S. & Laumann, K. (2020). Questionnaire measures and physiological correlates of presence: A systematic review. *Frontiers in Psychology, 11,* 1-21. https://doi.org/10.3389/fpsyg.2020.00349

Handel, S. (2019). *Perceptual organization: an integrated multisensory approach.* Springer. https://doi.org/10.1007/978-3-319-96337-2

Horikawa, T. & Kamitani, Y. (2017a). Generic decoding of seen and imagined objects using hierarchical visual features. *Nature Communications, 8*(15037), 1-15. https://doi.org/10.1038/ncomms15037

Horikawa, T. & Kamitani, Y. (2017b). Hierarchical neural representation of dreamed objects revealed by brain decoding with deep neural network features. *Frontiers in Computational Neuroscience, 11*(4), 1-11. https://doi.org/10.3389/fncom.2017.00004

Horikawa, T., Tamaki, M., Miyawaki, Y. & Kamitani, Y. (2013). Neural decoding of visual imagery during sleep. *Science, 340,* 639–642. https://doi.org/10.1126/science.1234330

Hume, D. (1739/2000). *A treatise of human nature.* Clarendon Press. https://doi.org/10.1093/oseo/instance.00046221

Husserl, E. (1975). *Experience and judgment.* Northwestern University Press.

Husserl, E. (2006). *Phantasy, image consciousness, and memory (1898-1925)* (Vol. 11). Springer Science & Business Media.

Ichikawa, J. (2008). Scepticism and the imagination model of dreaming. *The Philosophical Quarterly, 58*(232), 519–527. https://doi.org/10.1111/j.1467-9213.2007.546.x

Ichikawa, J. (2016). Imagination, dreaming, and hallucination. In *The Routledge Handbook of the Philosophy of Imagination* (pp. 149-162). Routledge.

Ichikawa, J. & Sosa, E. (2009). Dreaming, philosophical perspectives. In *The Oxford Companion to Consciousness.* Oxford University Press.

James, W. (1892/1962). *Psychology: briefer course.* Collier Books. https://doi.org/10.1037/11630-000

Jansen, J. A. (2010). Phenomenology, imagination and interdisciplinary research. In D. Schmicking & S. Gallagher (Eds.), *Handbook of Phenomenology and Cognitive Science* (pp. 141-158). Springer. https://doi.org/10.1007/978-90-481-2646-0_8

Kimble, M., Boxwala, M., Bean, W., Maletsky, K., Halper, J., Spollen, K. & Fleming, K. (2014). The impact of hypervigilance: evidence for a forward feedback loop. *Journal of anxiety disorders, 28*(2), 241-245. https://doi.org/10.1016/j.janxdis.2013.12.006

Kind, A. (2017). Imaginative vividness. *Journal of the American Philosophical Association, 3*(1), 32-50. https://doi.org/10.1017/apa.2017.10

LaBerge, S. (1981). Lucid dreaming: directing the action as it happens. *Psychology Today, 15*(1), 48-57.

LaBerge, S. (1985). *Lucid dreaming.* J. P. Tarcher.

LaBerge, S. (2000). Lucid dreaming: evidence and methodology. *Behavioral and Brain Sciences, 23*(6), 962-964. https://doi.org/10.1017/S0140525X00574020

LePort, A. K. R., Stark, S. M., McGaugh, J. L. & Stark, C. E. L. (2017). A cognitive assessment of highly superior autobiographical memory. *Memory, (Hove, English), 25*(2), 276-288. https://doi.org/10.1080/09658211.2016.1160126

Levy, L. (2012). Rethinking the Relationship between Memory and Imagination in Sartre's The Imaginary. *Journal of the British Society for Phenomenology, 43*(2), 143-160. https://doi.org/10.1080/00071773.2012.11006764

Loftus, E. F. & Pickrell, J. E. (1995). The formation of false memories. *Psychiatric annals, 25*(12), 720-725. https://doi.org/10.3928/0048-5713-19951201-07

Mahr, J. B. & Csibra, G. (2018). Why do we remember? The communicative function of episodic memory. *Behavioral and Brain Sciences, 41,* e1. https://doi.org/10.1017/S0140525X17000012

Mather, G. (2016). *Foundations of sensation and perception* (3 ed.). Psychology Press. https://doi.org/10.4324/9781315672236

Matthen, M. (2005). *Seeing, doing, and knowing: a philosophical theory of sense perception.* Oxford University Press. https://doi.org/10.1093/0199268509.001.0001

Matthen, M. (2010). Two visual systems and the feeling of presence. In N. Gangopadhyay, M. Madary & F. Spicer (Eds.), *Perception, action, and consciousness: sensorimotor dynamics and two visual systems* (pp. 107–124). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199551118.003.0007

McGurk, H. & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*(5588), 746-748. https://doi.org/10.1038/264746a0

Metzinger, T. (2003). Phenomenal transparency and cognitive self-reference. *Phenomenology and the Cognitive Sciences, 2*(4), 353-393. https://doi.org/10.1023/B:PHEN.0000007366.42918.eb

Metzinger, T. (2009). *The ego tunnel: the science of the mind and the myth of the self.* Basic Books.

Michaelian, K. (2011). Generative memory. *Philosophical Psychology, 24*(3), 323-342. https://doi.org/10.1080/09515089.2011.559623

Michaelian, K. (2016). Confabulating, misremembering, relearning: the simulation theory of memory and unsuccessful remembering. *Frontiers in Psychology 25*(1857), 1-13. https://doi.org/10.3389/fpsyg.2016.01857

Mitterer, H., Horschig, J. M., Müsseler, J. & Majid, A. (2009). The influence of memory on perception: it's not what things look like, it's what you call them. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(6), 1557. https://doi.org/10.1037/a0017019

Mizuguchi, N., Nakata, H., Uchida, Y. & Kanosue, K. (2012). Motor imagery and sport performance. *The Journal of Physical Fitness and Sports Medicine, 1*(1), 103-111. https://doi.org/10.7600/jpfsm.1.103

Estud.filos n.º 64. Julio-diciembre de 2021 | pp. 95-122 | Universidad de Antioquia | ISSN 0121-3628 | ISSN-e 2256-358X

119

https://doi.org/10.17533/udea.ef.n64a05

Morales, J. (2018). *The strength of the mind: essays on consciousness and introspection* (Ph. D. thesis). Columbia University.

Nanay, B. (2016a). Hallucination as mental imagery. *Journal of Consciousness Studies 23*(7-8), 65-81.

Nanay, B. (2016b). Imagination and perception. In *The Routledge handbook of the philosophy of imaginatio*n (pp. 124–134). Routledge.

Noë, A. (2004). *Action in perception.* The MIT Press.

Noë, A. (2007). Magic realism and the limits of intelligibility: what makes us conscious. *Philosophical Perspectives, 21,* 457-474. https://doi.org/10.1111/j.1520-8583.2007.00132.x

Overgaard, M. & Mogensen, J. (2017). An integrative view on consciousness and introspection. *Review of Philosophy and Psychology, 8,* 129-141. https://doi.org/10.1007/s13164-016-0303-6

Palombo, D. J., Sheldon, S. & Levine, B. (2018). Individual differences in autobiographical memory. *Trends in Cognitive Sciences, 22*(7), 583-597. https://doi.org/10.1016/j.tics.2018.04.007

Patihis, L., Frenda, S. J., LePort, A. K. R., Petersen, N., Nichols, R. M., Stark, C. E. L., McGaugh, J. L. & Loftus, E. F. (2013). False memories in highly superior autobiographical memory individuals. *Proceedings of the National Academy of Sciences, 110*(52), 20947-20952. https://doi.org/10.1073/pnas.1314373110

Pelaprat, E. & Cole, M. (2011). "Minding the gap": imagination, creativity and human cognition. *Integrative Psychological and Behavioral Science, 45*(4), 397-418. https://doi.org/10.1007/s12124-011-9176-5

Penfield, W. & Perot, P. (1963). The brain's record of auditory and visual experience. *Brain 86*(4), 595-696. https://doi.org/10.1093/brain/86.4.595

Perky, M. C. W. (1910). An experimental study of imagination. *The American Journal of Psychology, 21,* 422-452. https://doi.org/10.2307/1413350

Perrin, D., Michaelian, K. & Sant'Anna, A. (2020). The phenomenology of remembering is an epistemic feeling. *Frontiers in Psychology, 11*(1531), 1-14. https://doi.org/10.3389/fpsyg.2020.01531

Pizlo, Z. (2001). Perception viewed as an inverse problem. *Vision research, 41*(24), 3145-3161. https://doi.org/10.1016/S0042-6989(01)00173-0

Redshaw, J. (2014). Does metarepresentation make human mental time travel unique? *WIREs Cognitive Science, 5*(5), 519-531. https://doi.org/10.1002/wcs.1308

Revonsuo, A. (1995). Consciousness, dreams and virtual reality. *Philosophical Psychology, 8*(1), 35-58. https://doi.org/10.1080/09515089508573144

Robins, S. K. (2019). Confabulation and constructive memory. *Synthese, 196,* 2135-2151. https://doi.org/10.1007/s11229-017-1315-1

Rosen, M. G. (2013). What I make up when I wake up: anti-experience views and narrative fabrication of dreams. *Frontiers in Psychology*, 4. https://doi.org/10.3389/fpsyg.2013.00514

Rosen, M. G. (2018a). How bizarre? A pluralist approach to dream content. *Consciousness and Cognition, 62*, 148-162. https://doi.org/10.1016/j.concog.2018.03.009

Rosen, M. G. (2018b). Your dream-body: all an illusion? commentary on Windt's account of the dream-body. *Dreaming. Journal of Consciousness Studies, 25*(5-6), 44-62.

Rosen, M. G. (2019). Dreaming of a stable world: vision and action in sleep. *Synthese,* 1–36. https://doi.org/10.1007/s11229-019-02149-1

Russell, B. (1921). *The analysis of mind.* George Allen and Unwin.

Sanchez-Vives, M. V. & Slater, M. (2005). From presence to consciousness through virtual reality. *Nature Reviews Neuroscience, 6*, 332-339. https://doi.org/10.1038/nrn1651

Sant'Anna, A. & Michaelian, K. (2019). Thinking about events: a pragmatist account of the objects of episodic hypothetical thought. *Review of Philosophy and Psychology, 10,* 187-217. https://doi.org/10.1007/s13164-018-0391-6

Schacter, D. L. & Addis, D. R. (2007). The cognitive neuroscience of constructive memory: remembering the past and imagining the future. *Philosophical Transactions of the Royal Society B: Biological Sciences, 362*(1481), 773-786. https://doi.org/10.1098/rstb.2007.2087

Schredl, M. (2018). *Researching dreams: the fundamentals.* Palgrave Macmillan. https://doi.org/10.1007/978-3-319-95453-0

Schwitzgebel, E. (2008). The unreliability of naïve introspection. *Philosophical Review, 117*, 245-273. https://doi.org/10.1215/00318108-2007-037

Siegel, S. (2006). Subject and object in the contents of visual experience. *Philosophical Review, 115*(3), 355-88. https://doi.org/10.1215/00318108-2006-003

Sutton, J. (1998). *Philosophy and memory traces: Descartes to connectionism.* Cambridge University Press.

Sutton, J. (2003). Constructive memory and distributed cognition: towards an interdisciplinary framework. In B. Kokinov & W. Hirst (Eds.), *Constructive Memory* (pp. 290-303). New Bulgarian University.

Teroni, F. (2017). The phenomenology of memory. In *The Oxford handbook of philosophy of memory* (pp. 21–33). Oxford University Press. https://doi.org/10.4324/9781315687315-3

Thomas, N. J. (2014). The multidimensional spectrum of imagination: images, dreams, hallucinations, and active, imaginative perception. *Humanities, 3*(2), 132-184. https://doi.org/10.3390/h3020132

Tulving, E. (1983). *Elements of episodic memory.* Oxford University Press.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology/Psychologie Canadienne, 26*(1), 1–12. https://doi.org/10.1037/h0080017

Tulving, E. (2002). Episodic memory: from mind to brain. *Annual Review of Psychology, 53*, 1-25. https://doi.org/10.1146/annurev.psych.53.100901.135114

Voss, U., Schermelleh-Engel, K. Windt, J., Frenzel, C. & Hobson, A. (2013). Measuring consciousness in dreams: the lucidity and consciousness in dreams scale. *Consciousness and Cognition, 22*(1), 8-21. https://doi.org/10.1016/j.concog.2012.11.001

Wade, K. A., Garry, M., Read, J. D. & Lindsay, D. S. (2002). A picture is worth a thousand lies: using false photographs to create false childhood memories. *Psychonomic Bulletin & Review, 9*(3), 597-603. https://doi.org/10.3758/BF03196318

Whiteley, C. M. K. (2020). Aphantasia, imagination and dreaming. *Philosophical Studies, 178*, 2111–2132. https://doi.org/10.1007/s11098-020-01526-8

Wiese, W. & Metzinger, T. (2017). Vanilla PP for philosophers: a primer on predictive processing. In T. Metzinger & W. Wiese (Eds.), *Philosophy and Predictive Processing* (pp. 1-18). MIND Group. https://doi.org/10.7551/mitpress/9780262036993.003.0008

Windt, J. M. (2010). The immersive spatiotemporal hallucination model of dreaming. *Phenomenology and the Cognitive Sciences, 9*, 295-316. https://doi.org/10.1007/s11097-010-9163-1

Windt, J. M. (2018). Predictive brains, dreaming selves, sleeping bodies: how the analysis of dream movement can inform a theory of self- and world-simulation in dreams. *Synthese, 195*, 2577-2625. https://doi.org/10.1007/s11229-017-1525-6

A R T Í C U L O
D E I N V E S T I G A C I Ó N

Descartes y la memoria intelectual*

Diego Fabián Díaz Quiroz

Universidad Alberto Hurtado, Santiago, Chile
Email: diegodiazq92@gmail.com

Resumen: En el presente artículo se investiga la doctrina de la memoria intelectual en Descartes. En sus escritos, Descartes no solamente reconoció una memoria corporal, explicable en términos puramente fisiológicos, sino también una memoria intelectual o espiritual. En el presente artículo, se investiga si Descartes postuló una memoria intelectual por motivos teológicos o por motivos filosóficos. A partir del análisis de ciertos textos específicos en los cuales Descartes explica con relativo detalle en qué consiste la memoria intelectual, se intentará mostrar que Descartes, por motivos estrictamente filosóficos, apela a la memoria intelectual para explicar algunos procesos de reminiscencia que se producen en el ser humano. Las motivaciones de Descartes para postular una memoria intelectual no son teológicas, como han sostenido algunos comentadores contemporáneos de la doctrina cartesiana de la memoria.

Palabras clave: Descartes, memoria corporal, memoria intelectual, teología, filosofía

Cómo citar este artículo

Descartes and intellectual memory

Abstract: This article investigates the doctrine of intellectual memory in Descartes. In his writings, Descartes recognized not only a bodily memory, explainable in purely physiological terms, but also an intellectual or spiritual memory. In this article, I investigate whether Descartes postulated an intellectual memory for theological reasons or for philosophical reasons. From the analysis of certain texts in which Descartes explains what intellectual memory is, the paper will show that Descartes appeals, for strictly philosophical reasons, to intellectual memory to explain some processes of reminiscence that occur in the human being. In contrast to what some contemporary commentators on the Cartesian doctrine of memory have argued, Descartes's motivations for postulating intellectual memory are not theological.

Keywords: Descartes, corporeal memory, intellectual memory, theology, philosophy

Diego Fabián Diaz Quiroz

Licenciado en filosofía (Universidad Alberto Hurtado, Chile), Magíster en filosofía (Universidad Alberto Hurtado, Chile). Los títulos de las tesis de licenciatura y magíster son, respectivamente: *El rol de la imaginación en Descartes* (2015), y *La memoria intelectual en Descartes* (2019). Integrante de "Clio", grupo de estudios de historia de la filosofía, organizado por la Universidad Alberto Hurtado.

ORCID: 0000-0002-9377-742X

En el presente artículo se pretende investigar la motivación básica que llevó a Descartes a postular una memoria intelectual en algunos de sus escritos. El estudio de la memoria intelectual permite, entre otras cosas, aclarar aspectos semánticos y epistémicos de la filosofía de Descartes, así como su doctrina del intelecto (o "noética"). Se intentará mostrar que Descartes, por motivos estrictamente filosóficos, apela a una memoria intelectual para explicar algunos procesos de reminiscencia que se producen en el ser humano. Posteriormente, se argumentará en contra de la tesis según la cual Descartes ha postulado una memoria intelectual por motivos teológicos.

En primer lugar, se expondrá la motivación teológica que llevó a algunos pensadores del medioevo y la modernidad temprana a postular una memoria intelectual. Posteriormente, se revisarán los textos en los cuales Descartes explica su doctrina de la memoria intelectual, comenzando por la conversación entre Descartes y el autor de apellido Burman,[1] para continuar con la correspondencia entre Descartes y Antoine Arnauld (1612–1694). Finalmente, se discutirá la tesis según la cual Descartes ha postulado una memoria intelectual por motivos teológicos.

1. Concepciones medievales y modernas de la memoria intelectual con motivaciones teológicas

En algunos intelectuales del medioevo, la memoria intelectual es postulada por motivos fundamentalmente teológicos. En efecto, en la *Suma teológica*, Tomás de Aquino sostiene que incluso las almas "separadas" están dotadas de memoria, apelando directamente a las sagradas escrituras: "la memoria permanece en el alma separada [*anima separata*], pues se dice en Lc 16, 25, al rico Epulón, cuya alma estaba en el infierno: *Acuérdate que recibiste bienes mientras vivías.* Por lo tanto, en el alma separada permanece la memoria" (I, 77, art.8).[2] Si las escrituras nos enseñan que incluso después de la muerte el alma podrá recordar las experiencias de su existencia previa, entonces el alma, ya separada del cuerpo, debe poseer una memoria que le permita acceder a sus experiencias del pasado.

Siguiendo esta tendencia, algunos autores de la modernidad temprana también han postulado una memoria intelectual por motivos teológicos. El autor de apellido Burman, en su entrevista con Descartes, también postula una memoria intelectual, no solo en las almas separadas, sino también en los ángeles:

---

1   No se sabe con exactitud quién es este autor de apellido Burman. Sobre la historia y valor interpretativo de esta conversación, véase Beyssade (2001).

2   Esta traducción al español es de la edición Leonina de las *opera omnia* de Santo Tomás de Aquino (1889). Todas las traducciones al español de este trabajo son propias.

Estud.filos n.º 64. Julio-diciembre de 2021 | pp. 123-138 | Universidad de Antioquia | ISSN 0121-3628 | ISSN-e 2256-358X

125

https://doi.org/10.17533/udea.ef.n64a06

> Pero aunque no se impriman vestigios en el cerebro, y de este modo no haya memoria corporal, existe [*datur*], sin embargo, la memoria intelectual [*memoria intellectualis*], como en los ángeles y en las almas separadas [*animabus separatis*], sin duda, y así, mediante ella, la mente recordaría sus pensamientos (AT V, 150).[3]

Es comprensible pensar, siguiendo a un autor contemporáneo, que Burman está siguiendo la misma motivación teológica que Tomás de Aquino (Joyce, 1997). ¿Sucede lo mismo con Descartes? Después de todo, Descartes podría conceder perfectamente el argumento ofrecido por Tomás de Aquino acerca de la memoria intelectual.[4] Como señala Descartes en su carta a Mersenne, fechada el 15 de abril de 1630, todo "lo que depende de la revelación" (AT I, 144) pertenece a la teología. La revelación teológica proviene de la "luz de la gracia [*lumen gratiae*]" (AT VII, 148) o "luz de la fe [*lumière de la Foy*]" (AT V, 82), que pertenece a Dios mismo, y que asegura a los fieles la certidumbre absoluta de lo revelado por Dios. Ahora bien, según los planteamientos de Descartes en las *Meditaciones metafísicas*, es Dios quien "da la gracia" (AT VII, 2) para que los fieles crean no solo en su existencia, sino también en las "demás cosas" (AT VII, 2). Por consiguiente, "se debe creer en las sagradas escrituras" (AT VII, 2), las cuales "son tomadas de Dios" (AT VII, 2). Por lo tanto, una memoria perteneciente a las almas separadas sería completamente admisible, con la ayuda fundamental de las sagradas escrituras y la "luz de la fe" otorgada por Dios. Si esto es así, entonces Descartes, siguiendo a los autores previamente mencionados, podría postular una memoria intelectual por motivos exclusivamente teológicos; después de todo, como señalan algunos autores contemporáneos, "[l]a escatología siempre había sido el sentido [*point*] de las doctrinas de la memoria intelectual" (Sutton, 2007, p. 69).

Veamos, en lo que sigue, cuáles son las motivaciones que tuvo Descartes para postular una memoria intelectual.

2. La conversación entre Descartes y Burman acerca de la memoria intelectual

En la conversación fechada el 16 de abril de 1648, Burman sostiene que la mente puede pensar muchas cosas a la vez (AT V, 150). Descartes concede esta afirmación de Burman, declarando, además, que en ocasiones no recordamos los pensamientos de nuestra juventud, porque en el cerebro no ha quedado impreso "ningún vestigio

---

3   Se ha utilizado la edición de C. Adam y P. Tannery para citar las obras de Descartes (1897 – 1913).

4   No hay duda de que Descartes tenía al menos una obra de Tomás de Aquino en su poder. En su carta a Mersenne, fechada el 25 de diciembre de 1639, Descartes declara que posee "una Suma de S. Tomás" (AT II, 630). Sin embargo, Descartes no dice en su carta qué obra es en específico.

[*vestigia*]" (AT V, 150) de ellos. En efecto, según Descartes, los pensamientos que ha tenido nuestra mente a lo largo de la vida dejan, en ocasiones, ciertas huellas o vestigios en el cerebro, que nos permiten recordar esos pensamientos. Muchos de esos vestigios fisiológicos, o como los llama Descartes en el *Tratado del hombre,* "impresiones de la memoria" (AT XI, 184), quedan almacenados principalmente en cierta zona del cerebro, y permiten, posteriormente, la activación de mecanismos fisiológicos que producen determinados efectos en los espíritus animales que transitan sobre la superficie de la glándula pineal. Este proceso, a su vez, afecta a la mente de tal manera, que la mente vuelve a tener los mismos pensamientos que había tenido en el pasado, aunque el pensamiento en cuestión puede tener un grado de vivacidad o "claridad" distinto. Por ejemplo, si alguna vez observé algún objeto físico individual del mundo, dicho objeto causa en mi cerebro una serie de efectos, que terminan afectando a los espíritus animales que transitan sobre la superficie de la glándula pineal, disponiéndolos de una forma determinada. A la disposición específica de esos espíritus que se desplazan sobre la glándula, Descartes la denomina en el *Tratado* con el nombre de *idea* (*idée,* AT XI, 176-177). Cuando la mente "advierte" esa idea, la mente se ve afectada por una sensación (en este caso, se trata de una sensación visual, causada por el objeto observado). Ahora bien, en algunos casos, no solo se produce la idea en la glándula pineal, sino también una serie de "figuras" (*figures*) en otras zonas del cerebro. Estas figuras (es decir, ciertos pliegues o disposiciones de las fibras nerviosas del cerebro) permiten posteriormente formar la misma idea en la glándula pineal que había sido formada por el objeto observado en el pasado, sin la necesidad de la presencia efectiva de dicho objeto externo. De este modo, la mente vuelve a "contemplar" la idea del objeto que había observado en el pasado, y así puede recordarlo.[5] Sin las huellas materiales de la memoria, sería imposible recordar el objeto visto previamente. Por esta razón, Descartes declara en su conversación con Burman que, debido a la ausencia de los vestigios fisiológicos necesarios, no recordamos los pensamientos que tuvimos en diversas etapas pasadas de la vida.

A continuación, Burman menciona brevemente el tema de la memoria intelectual, afirmando lo siguiente:

> Pero aunque no se impriman vestigios en el cerebro, y de este modo no haya memoria corporal, existe [*datur*], sin embargo, la memoria intelectual [*memoria intellectualis*], como en los ángeles y en las almas separadas [*animabus separatis*], sin duda, y así, mediante ella, la mente recordaría sus pensamientos (AT V, 150).

---

5   Para una explicación más detallada de este proceso de reminiscencia, véase el *Tratado del hombre* (AT XI, 177 y ss.). Algunos planteamientos del *Tratado* vuelven a aparecer en el *Discurso del método* y *Las pasiones del alma,* prácticamente de la misma manera.

Aquí Burman afirma brevemente que existe una memoria completamente inmaterial, mediante la cual la mente sería capaz de recordar sus pensamientos. En este pasaje, Burman sugiere que esta memoria es absolutamente independiente del cuerpo, puesto que ella podría darse sin los vestigios fisiológicos del cerebro. Según Burman, los ángeles y las "almas separadas" estarían dotados de esta memoria.

Descartes responde lo siguiente:

> No niego la memoria intelectual [*memoriam intellectualem*]; en efecto, ella existe [*datur*]. Como cuando, al escuchar que la palabra R-E-Y significa poder supremo, aprendo aquello de memoria, y sucesivamente evoco, mediante la memoria, aquel significado. Aquello, ciertamente, se produce por la memoria intelectual, puesto que no hay ninguna afinidad [*affinitas*] entre aquellas tres letras y su significado a partir de la cual extraerlo, sino que, por medio de la memoria intelectual, recuerdo lo que esas letras denotan. Pero esta memoria intelectual es más de universales que de particulares; y de este modo, mediante ella no podemos recordar todas nuestras acciones particulares (AT V, 150).

Para articular su respuesta, Descartes no hace referencia ni a los ángeles ni a las almas separadas, sino al aprendizaje del significado de las palabras. En efecto, es imposible, argumenta Descartes, extraer el significado de una palabra a partir de algún tipo de afinidad o parentesco entre las letras de esa palabra y su significado, puesto que tal afinidad es inexistente. Según lo que Descartes plantea en este pasaje, lo único que puede hacerse es aprender el significado de la palabra *rey* y después recordar, mediante la memoria intelectual, aquel significado. La memoria corporal jamás podría proporcionarnos, según Descartes, dicho recuerdo.[6]

---

6   Ya en las *Reglas para la dirección del espíritu*, Descartes distinguía entre memoria corporal e intelectual, aunque sutilmente. En efecto, Descartes sostiene en las *Reglas* lo siguiente: "pero aquella memoria [*memoria*], al menos la que es corpórea y semejante a la memoria de los brutos [*saltem quae corporea est & similis recordationi brutorum*], no es en nada distinta de la imaginación" (AT X, 416). Aquí Descartes nos habla de una memoria corporal que tanto seres humanos como otros animales poseen. *Al menos* esta memoria, según Descartes, podemos encontrarla en seres humanos y otros animales. Al decir esto, Descartes parece estar suponiendo que hay otro tipo de memoria, además de la corporal. Dado que Descartes probablemente leyó, cuando estudiaba en *La Flèche*, los comentarios a los *Parva naturalia* de Aristóteles realizados por los jesuitas de Coímbra, en la obra titulada *Comentarii collegii Conimbricensis societatis Iesu in libros Aristotelis, qui Parva Naturalia appellantur* (1596), es probable que haya encontrado allí por primera vez la distinción entre dos tipos de memoria. Durante la década de 1640, Descartes intenta entrar en contacto nuevamente con los jesuitas (después de veinte años). En una carta a Mersenne, fechada el 3 de diciembre de 1640, Descartes reconoce haber leído la filosofía de Charles François d'Abra de Raconis (probablemente, Descartes leyó el *Totius philosophiae, hoc est Logicae, Moralis, Physicae, et Metaphysicae, brevis et accurata, facilique et clara methodo disposita tractatio*), los comentarios de los jesuitas de Coímbra, y a Eustaquio de San Pablo, cuya obra más importante y difundida en la época era la *Summa philosophicae quadripartita de rebus dialecticis, moralibus, physicis, et metaphysicis*, publicada por primera vez en 1609 (AT III, 251). Esta última obra fue muy elogiada por Descartes en una carta a Mersenne, fechada el 11 de noviembre de 1640 (AT III, 232). En los *Comentarios*, los jesuitas de Coímbra relacionaban la memoria intelectiva con el recuerdo de las cosas universales e inmateriales (Coímbranses, 1596), y discutían si acaso la memoria intelectiva y el intelecto eran facultades distintas. Descartes no niega que la memoria intelectual pueda recordar particulares, aunque señala que ella es "más de universales" (AT V, 150). Por su parte, en la *Summa* de Eustaquio

Ahora bien, ¿por qué apelar a una memoria intelectual en este contexto? ¿Por qué no apelar simplemente al funcionamiento de la memoria corporal para explicar este caso de reminiscencia? Algunas observaciones relativamente breves respecto a las palabras o signos convencionales podrían permitirnos comprender por qué Descartes considera necesario atribuir a una memoria que no es corporal el recuerdo del significado de las palabras.

Como se ha mencionado más arriba, lo único que puede retener la memoria corporal son diversas impresiones sensibles, que pueden ser causadas, por ejemplo, por diversos cuerpos del mundo exterior, que afectan a los órganos de nuestros sentidos de diversas maneras. La memoria corporal puede realizar aquel proceso de almacenamiento gracias a los vestigios fisiológicos implantados en el organismo, principalmente en el cerebro. Ahora bien, en una carta de agosto de 1641 a Hyperaspistes, Descartes declara que las palabras, o más específicamente, los nombres (*nomina*) "son corpóreos" (AT III, 425). De manera similar, aunque con un poco más de detalle y claridad, Descartes escribe a Chanut, en una carta fechada el 1 de febrero de 1647, lo siguiente: "[c]uando se aprende una lengua [*langue*], se unen las letras o la pronunciación de ciertas palabras, que son cosas materiales [*choses materielles*], con sus significados, que son pensamientos" (AT IV, 604). Por otro lado, Descartes, en *Las pasiones del alma*, sostiene que las palabras (*paroles*) "según la institución de la naturaleza, solo representan al alma su sonido, cuando son proferidas por la voz, o la figura de sus letras, cuando son escritas" (AT XI, 369). Tanto en este último pasaje como en la correspondencia citada un poco más arriba, podemos reconocer los rasgos físico-sensibles básicos de las palabras identificados por Descartes. La memoria corporal solo puede retener las impresiones sensibles que las palabras pueden causar en el cerebro, como los sonidos de sus sílabas cuando son pronunciadas, o las figuras de sus letras cuando son escritas. Por consiguiente, si utilizo mi memoria corporal, lo único que puede proveer aquella memoria es el recuerdo de los rasgos sensibles de las palabras, como el sonido de sus sílabas o la figura de sus letras, puesto que las palabras (escritas o pronunciadas) son "cosas materiales". En consecuencia, la memoria corporal no puede, por sí sola, ayudar a la mente a recordar qué es lo que esas palabras significan. Siguiendo lo planteado por Descartes en su carta a Chanut, entender el significado de las palabras es entender a qué pensamientos ellas suelen estar unidas por convención. Para recordar esto último, es necesario apelar a otro tipo de memoria que no es corporal, y que reside en la mente.[7]

---

de San Pablo se discute muy brevemente el tema de la memoria. Eustaquio de San Pablo sostiene que la memoria no es una facultad distinta de la fantasía (*phantasia*, Eustanquio de San Pablo, 1647), y establece una distinción entre memoria (*memoria*) y reminiscencia (*reminiscentia*). La segunda solo el ser humano la posee, debido a su "capacidad discursiva" (*vis discursus*, Eustaquio de San Pablo, 1647, p. 273). La reminiscencia sería posible gracias al raciocinio (*ratiocinatio*), que está ausente en el resto de los animales. Descartes utiliza los verbos *memini* y *recordari*, pero no hace una distinción entre ellos. Para un examen detallado del contexto histórico-filosófico de Descartes, véase el excelente trabajo de Roger Ariew (2014).

7   Entre los diversos diccionarios filosóficos disponibles de Descartes publicados por autores especializados, el de John Cottingham (1993) es quizá el único que ha incluido un breve análisis del papel que juega la memoria intelectual en el

Algunos autores han criticado el argumento que Descartes ofrece a Burman. Por ejemplo, Joyce (1997) considera el argumento de Descartes "extraño" (p. 385), ya que Descartes a menudo hace referencia a la "habilidad de los animales para asociar dos cosas distintas y arbitrariamente conectadas, y esto [Descartes] lo suponía explicable en términos puramente corpóreos" (Joyce, 1997, p. 385). Descartes, en efecto, admitió procesos de aprendizaje asociativo en algunos autómatas o "animales-máquina".[8] Por lo tanto, argumenta Joyce (1997), un autómata podría "asociar, vía ostensión, un sonido con un objeto" (p. 386); a su vez, tal autómata estaría en condiciones de aprender asociaciones cada vez más complejas entre sonidos y objetos. Por consiguiente, en el aprendizaje de las palabras no hay nada tan sorprendente como para postular una memoria intelectual.

Lo que parece sugerir Joyce es que un autómata podría asociar, mediante gestos ostensivos, un sonido y una cosa, que estarían vinculados arbitrariamente. Ahora bien, el argumento de Descartes apunta en otra dirección; no se trata meramente de casos de "referencialidad", es decir, de casos en los cuales se asocia una palabra con una cosa que existe fuera del pensamiento, sino de una evaluación semántica o de significado, es decir, de una vinculación entre pensamientos y palabras. Los animales no pueden saber a qué pensamientos están ligadas las palabras de un lenguaje determinado. En efecto, un animal no puede saber en qué estamos pensando cuando usamos la palabra *rey* (de hecho, ni siquiera todo ser humano lo sabe; solo lo sabe el que conoce el idioma español). La relación, más que entre palabras y cosas, es entre palabras e ideas o conceptos de nuestra mente. En el argumento presentado a Burman, Descartes afirma que la memoria intelectual es "más de universales". Como señala Descartes en *Los principios de la filosofía*, los universales solo son modos del pensar (AT VIII, 27); no tienen, por así decir, realidad extramental. Un animal, por ejemplo, podría pronunciar la palabra *triángulo*, pero no puede comprender en qué pensamos nosotros cuando pronunciamos o escuchamos ese nombre universal. La palabra *triángulo* es un nombre universal, que remite, por tanto, a una idea universal de nuestra mente. Lo que debemos hacer, si queremos recordar el significado de aquel nombre universal, es recordar a qué idea dicho nombre ha sido vinculada por convención. Es aquí donde interviene la memoria intelectual.[9]

---

aprendizaje del lenguaje. Por su parte, en el diccionario de Frédéric de Buzon & Denis Kambouchner (2011), así como en el de Kurt Smith (2015), no hay ninguna referencia a la memoria. Finalmente, en el diccionario de Ariew, Des Chene, Jesseph, Schmaltz, y Verbeek (2015) no se ha destacado la relevancia de la memoria intelectual en el aprendizaje del lenguaje.

8   Véase, *v. gr.*, *Las pasiones del alma*, primera parte, artículo L (AT XI, 368).

9   En la primera parte de *Los principios de la filosofía*, Descartes dedica una sección a las causas principales de nuestros errores. La cuarta causa, según el artículo LXXIV, es que ligamos (*alligemus*) nuestros conceptos a palabras que no corresponden cuidadosamente a las cosas. En este artículo, Descartes afirma lo siguiente: "por el uso del lenguaje, ligamos [*alligamus*] todos nuestros conceptos a palabras con las cuales los expresamos, y no los aprendemos de memoria a no ser que los aprendamos junto con estas palabras [*nec eos nisi simul cum istis verbis memoriae mandamus*]" (AT VIII, 37). Según este pasaje de *Los Principios*, durante nuestro aprendizaje, ligamos, mediante la memoria, palabras e ideas (o conceptos). Descartes debería

3. La correspondencia entre Descartes y Antoine Arnauld

En la correspondencia entre Descartes y Arnauld podemos encontrar nuevamente un caso de reminiscencia en el cual la intervención de una memoria que no se encuentra en ningún órgano del cuerpo es, según Descartes, condición necesaria para que aquella reminiscencia pueda darse. Veamos lo que plantea Descartes al respecto.

En su carta del 3 de junio de 1648, Arnauld escribe algunas observaciones respecto a la física y metafísica propuesta por Descartes. Con respecto a la metafísica, Arnauld presenta en su carta algunas inquietudes sobre Dios y la mente humana. Una de las inquietudes de Arnauld sobre la mente humana se relaciona precisamente con la memoria. En efecto, a Arnauld le sorprende que Descartes haya afirmado, por una parte, que la mente siempre esté pensando, y que, por otra, no podamos recordar ningún pensamiento que hayamos tenido en el útero materno o en estados letárgicos. Después de manifestar esta inquietud, Arnauld propone lo siguiente: "parece que en nuestra mente hay que admitir necesariamente un doble poder de la memoria [*duplex memoriae vis*]: una puramente espiritual [*mere spiritualis*], y otra que necesita [*indigeat*] de algún órgano corpóreo" (AT V, 186).

Al igual que Arnauld, Descartes reconoce dos tipos de memoria en los seres humanos: "admito un doble poder de la memoria [*duplicem memoriae vim*]" (AT V, 192). Además, Descartes, en esta primera respuesta a Arnauld, explica brevemente en qué consistiría esta memoria intelectual. Sin embargo, en primer lugar, Descartes asegura estar persuadido de que en la mente del infante "nunca hubo intelecciones puras, sino solamente sensaciones confusas" (AT V, 192). Además, según Descartes, por más que queden vestigios fisiológicos en el cerebro de esas sensaciones, ellos

> no bastan [*sufficiunt*] para advertir [*advertamus*] que las sensaciones que nos llegan cuando somos adultos son similares a las que tuvimos en el útero materno, y así para que las recordemos; porque esto depende [*pendet*] de una reflexión del intelecto o memoria intelectual [*reflexione intellectus sive memoriae intellectualis*], la cual no era de ninguna utilidad en el útero (AT V, 192-93).

Aquí podemos observar una primera aproximación de Descartes en su correspondencia con Arnauld a la memoria intelectual. En este caso, el recuerdo se produciría gracias a una comparación entre una sensación presente y las sensaciones del pasado que han quedado "registradas" en el cerebro, por medio de vestigios fisiológicos. La memoria intelectual consistiría en una reflexión mediante la cual nos damos cuenta de que una sensación que experimentamos en el presente es similar a una sensación del pasado.

---

admitir, a pesar de que no lo haya afirmado explícitamente, que en este caso no solo es necesaria la intervención de la memoria corporal, sino también la intervención de la memoria intelectual. En efecto, solo ella nos permite recordar a qué pensamientos están ligadas nuestras palabras.

No es casual que Descartes considere este tipo de comparación como un caso genuino de memoria, puesto que, como afirma Descartes en las *Meditaciones*, una de las características generales de la memoria es que ella "conecta [*connectit*] lo presente con lo precedente" (AT VII, 89). Gracias a la memoria intelectual, podemos conectar (con o sin éxito) sensaciones presentes con sensaciones pasadas almacenadas mediante vestigios fisiológicos cerebrales.

En una segunda carta de Descartes a Arnauld, y continuando con el tema de la memoria intelectual, Descartes aclara un poco más sus planteamientos sobre este asunto. Descartes afirma lo siguiente:

> No es suficiente [*satis*], para recordar alguna cosa, que aquella cosa se haya presentado [*obversata sit*] antes a nuestra mente, y que haya dejado algunos vestigios en el cerebro, los cuales dan la ocasión para que la misma [cosa] se presente [*occurrit*] nuevamente a nuestro pensamiento, sino que, además, se requiere [*requiritur*] que reconozcamos [*agnoscamus*], cuando [esa cosa] se presenta por segunda vez, que esto sucede porque antes había sido percibida por nosotros (AT V, 219-220).

En este pasaje, Descartes argumenta que, para recordar alguna cosa, deben cumplirse tres requisitos: (i) que esa cosa se haya presentado al pensamiento antes, al menos una vez; (ii) que esa cosa deje un vestigio fisiológico en el cerebro que la represente; (iii) que podamos reconocer, cuando esa cosa vuelve a presentarse al pensamiento, que ella se está presentando por segunda vez. Siguiendo la primera respuesta de Descartes a Arnauld, este reconocimiento sólo es atribuible a la reflexión realizada por el intelecto o entendimiento, que Descartes denomina en este contexto *memoria intelectual*. En efecto, las impresiones sensibles (sea cual sea su origen y su grado de vivacidad) solo constituyen causas ocasionales que afectan a la mente de diversas maneras. No obstante, ninguna impresión sensible, por sí misma, es capaz de "comunicar" a la mente su novedad. Para detectar la novedad de una impresión sensible que afecta a la mente, se requiere un examen del intelecto, mediante el cual este último compara las impresiones sensibles que el cuerpo le ofrece. Descartes afirma que, para reconocer la novedad de una impresión sensible, la mente "debió utilizar una intelección pura" (AT V, 220), ya que "no puede haber ningún vestigio corporal de esta novedad" (AT V, 220). Descartes da el siguiente ejemplo: "a menudo a los poetas se les presentan [*occurrunt*] algunos versos, los cuales no recuerdan haber leído alguna vez junto a otros, versos que, sin embargo, no se les presentarían, si no los hubiesen leído en otro lugar" (AT V, 220). Si consideramos este ejemplo desde una perspectiva en "tercera persona", podríamos decir que los poetas efectivamente "recuerdan" los versos, ya que, gracias a sus vestigios fisiológicos cerebrales, los versos que habían leído vuelven a presentarse a sus pensamientos. No obstante (y es lo que intenta poner de relieve Descartes), y desde una perspectiva en primera persona, los poetas necesitan reflexionar para

percibir, mediante su intelecto, que los versos que se les presentan al pensamiento ya los habían leído antes. Según Descartes, ninguna huella fisiológica rememorativa podría causar ese efecto en la mente de los poetas. Los vestigios solo causan en la mente el mismo pensamiento, pero no pueden comunicar nada acerca de su novedad. Para que suceda esto último, se requiere una especie de *inspectio mentis*, reflexión del intelecto, o memoria intelectual. Finalmente, Descartes realiza una analogía para explicar el tipo peculiar de reflexión de la memoria intelectual:

> así como distinguimos entre la visión directa y la refleja [*reflexam*], en que aquella depende del primer encuentro de los rayos de luz, [y] esta [depende] del segundo, los pensamientos primeros y simples de los infantes (cuando sienten dolor […] o placer [*voluptas*] […]) los llamo *directos*, no reflexivos; cuando, en cambio, un adulto siente algo, y junto con ello percibe [*percipit*] que no ha sentido lo mismo antes, a esta segunda percepción la llamo *reflexión*, y la remito solo al intelecto [*ad intellectum solum refero*], aunque haya estado unida de tal manera a la sensación, que se producen juntas [*simul*], y no parecen distinguirse una de otra (AT V, 220-21).

Así como la visión directa (*visio directa*) depende del primer encuentro con los rayos de luz, los pensamientos directos de los infantes (*cogitationes directae*), como el dolor y el placer, dependen del primer encuentro del alma con algo que, en este caso, es ajeno a ella (los procesos fisiológicos del cuerpo), y que la afecta de cierta manera. Por su parte, así como la visión refleja (*visio reflexa*) depende del segundo encuentro con los rayos de luz, la reflexión del intelecto que realiza un adulto depende de que se le presente al pensamiento algo por segunda vez. Según el ejemplo de Descartes en el pasaje que se acaba de citar, se presenta algo mediante los sentidos que nunca había sido percibido. A pesar de que lo percibido en este caso es nuevo, no podría ser percibido como tal si no tuviéramos alguna sensación pasada (registrada en el cerebro por medio de vestigios fisiológicos) con la cual compararlo. Por esta razón, se requiere que algo ya se nos haya presentado al pensamiento mediante los sentidos.

4. ¿Motivaciones teológicas?

Algunos autores sostienen que Descartes ha postulado una memoria intelectual por motivos fundamentalmente teológicos. John Sutton (2007), por ejemplo, después de analizar la respuesta de Descartes a Burman, y después de sostener que la memoria corporal asociativa hubiese sido suficiente para explicar el aprendizaje del significado de las palabras, afirma, con respecto a Descartes, que "el contexto clave para sus discusiones de la memoria intelectual es aquel, bastante diferente, de la supervivencia y la inmortalidad" (p. 69). Además, Sutton (2007) declara que "[l]a escatología siempre había sido el sentido [*point*] de las doctrinas de la memoria intelectual" (p. 69). Según

Estud.filos n.º 64. Julio-diciembre de 2021 | pp. 123-138 | Universidad de Antioquia | ISSN 0121-3628 | ISSN-e 2256-358X

133

https://doi.org/10.17533/udea.ef.n64a06

Sutton, algunos autores escolásticos pensaban que era necesario postular una memoria intelectual para que las "almas separadas" o "desencarnadas" (*disembodied*) pudiesen recordar el pasado después de la muerte (Sutton, 2007).[10] Descartes habría postulado la memoria intelectual fundamentalmente para continuar dentro de este marco contextual histórico preexistente.

Dada la evidencia textual, esta tesis no parece demasiado plausible. En primer lugar, y como se ha mostrado más arriba, Descartes, en sus explicaciones sobre la memoria intelectual, no hace ninguna referencia al tema de la inmortalidad y supervivencia del alma después de la muerte. Estos asuntos acerca de la existencia *post mortem* no tienen ninguna relevancia en la argumentación de Descartes en este contexto. En segundo lugar, la memoria intelectual de la que habla Descartes en sus cartas a Burman y Arnauld depende, en cierta medida, del cuerpo. En efecto, el intelecto necesita de las huellas fisiológicas del cerebro para compararlas y realizar su acción reflexiva; es imposible reflexionar acerca de la novedad de una sensación presente sin compararla con sensaciones pasadas almacenadas en el cerebro por medio de impresiones fisiológicas de la memoria. En cuanto al rol de la memoria intelectual en el aprendizaje del lenguaje, el entendimiento requiere de los nombres cuyo significado debe recordar, y los nombres, según Descartes, son corpóreos, o cosas materiales. Por consiguiente, la memoria corporal también es necesaria; sin ella, sería imposible recordar los rasgos físico-sensibles de las palabras. En este caso, así como en el caso de la "memoria intelectual reflexiva", se trata de recuerdos que un "alma separada" o "desencarnada" no podría tener (el que habla de almas separadas no es Descartes, sino Burman). Los casos de memoria intelectual analizados por Descartes son procesos de reminiscencia que se producen en seres humanos, es decir, en entidades sustancialmente compuestas, y tales procesos son inexplicables si no apelamos, por razones filosóficas, a una memoria que no reside en el cuerpo.

Por otro lado, y en tercer lugar, es difícil creer que Descartes aceptaría, sin más, alguna creencia ortodoxa teológica preexistente. Como señala Joyce (1997), Descartes, en algunos casos, estaba dispuesto a "desafiar a la iglesia cuando su método filosófico lo condujo claramente a opiniones contrarias" (p. 382). Un ejemplo de esta actitud desafiante podría ser la tesis cartesiana según la cual Dios no puede ser engañador. Como es sabido, y como podemos observar en las segundas objeciones contra las *Meditaciones* de Descartes recogidas por Mersenne, algunos teólogos de la época se oponían a esta tesis (AT VII, 125).[11] Sin embargo, Descartes mantuvo su posición

---

10   Sutton no menciona ningún autor en particular.

11   Según el argumento de estos teólogos, recogido por Mersenne en las segundas objeciones, Dios puede mentir, ya que, en la biblia, Dios dice muchas cosas que no sucedieron. Descartes responde que su posición filosófica es compatible con la teología, ya que se basa en la distinción (comúnmente aceptada entre los teólogos, según Descartes) entre dos maneras de hablar de Dios (*modi loquendi de Deo*): una "acomodada al vulgo" (AT VII, 142), y por tanto siempre relacionada a los hombres, y otra sin relación a los hombres, que expresa una "verdad más desnuda" (AT VII, 142). A juicio de Descartes, esta

respecto a este tema, a pesar de esa creencia teológica aceptada por algunos teólogos de la época.

Ahora bien, es importante reconocer lo siguiente. En una oportunidad, Descartes efectivamente habla de una memoria intelectual en un "contexto teológico", a saber, en una carta a Huygens del 13 de octubre de 1642. Allí, Descartes, tratando de consolar a su amigo, declara que podremos recordar a nuestros seres queridos después de la muerte, "pues encuentro en nosotros una memoria intelectual, que es ciertamente independiente del cuerpo" (AT III, 580). No obstante, esta memoria intelectual no es la que explica Descartes en sus cartas a Burman y Arnauld, puesto que se trata de una memoria totalmente independiente del cuerpo. En cambio, la memoria intelectual que postula Descartes en sus cartas a Burman y Arnauld depende, al menos en parte, del cuerpo, ya que el intelecto debe aplicar sus capacidades a impresiones sensibles para acceder a los recuerdos requeridos. Ahora bien, incluso en este caso, no hay por qué pensar que Descartes esté postulando una memoria intelectual simplemente para continuar con una tradición ortodoxa escatológica, puesto que Descartes había hecho algunas breves referencias anteriores a esta memoria en contextos no teológicos, antes de su carta a Huygens. En efecto, en una carta del 1 de abril de 1640, y después de realizar algunas observaciones con respecto a la memoria corporal, Descartes escribe a Mersenne lo siguiente: "[p]ero, además de esta memoria, que depende del cuerpo, aún reconozco otra, totalmente intelectual [*du tout intellectuelle*], que solo depende del alma [*qui ne dépend que de l'ame seule*]" (AT III, 48). En otra carta dirigida a Mersenne del 4 de junio de 1640, y sin ninguna referencia a la escatología, Descartes sostiene que "la memoria intelectual tiene sus especies aparte, que no dependen de ningún modo de esos pliegues [cerebrales]" (AT III, 84-85). En estas cartas a Mersenne, Descartes declara que la mente, por sí sola, cuenta con una capacidad rememorativa, y que el alma posee sus propios vestigios, o "especies" (*especes*) inmateriales. En una carta a Mesland del 2 de mayo de 1644, y nuevamente en un contexto no teológico, Descartes afirma que la memoria de las cosas materiales (*choses materielles*) depende de los vestigios que permanecen en el cerebro, pero la memoria de las cosas intelectuales (*choses intellectuelles*) "depende de otros vestigios, que permanecen en el pensamiento mismo" (AT IV, 114). Sin embargo, se presentan dificultades considerables para examinar esta memoria puramente intelectual, ya que las referencias que hace Descartes son muy breves. En ningún lugar Descartes nos explica en qué consiste esta memoria que solo depende del espíritu, y cuál es el funcionamiento de los vestigios mentales que

---

última manera de hablar es la que debe usarse cuando se filosofa. Si en la escritura hubiese algún lugar donde se encuentra una mentira de Dios, no sería más que una mentira "que se expresa con palabras" (AT VII, 143), mientras que Descartes, en sus *Meditaciones*, no ha hablado de esa mentira meramente verbal, sino de la "malicia interna y formal que está contenida en el engaño" (AT VII, 143). Además, Descartes refuerza su contraargumento, señalando que en los pasajes aludidos por los teólogos Dios no está mintiendo; solo está amenazando o negando beneficios.

Estud.filos n.º 64. Julio-diciembre de 2021 | pp. 123-138 | Universidad de Antioquia | ISSN 0121-3628 | ISSN-e 2256-358X

135

https://doi.org/10.17533/udea.ef.n64a06

la hacen posible. En la carta a Mesland que se acaba de citar, Descartes se limita a decir de los vestigios intelectuales lo siguiente: "no sabría explicarlos por algún ejemplo sacado de las cosas corporales, que no sea muy diferente" (AT IV, 114). En esta oportunidad, Descartes reconoce su incapacidad para encontrar algún objeto o proceso físico que nos permita comprender el funcionamiento de las huellas espirituales rememorativas.[12]

Así, Descartes parece haber distinguido, por razones meramente filosóficas, al menos tres tipos de memoria: (i) memoria corporal (ya sea en el cerebro o en el resto del cuerpo); (ii) memoria intelectual (cuya explicación podemos encontrar en la correspondencia con Arnauld y en la conversación con Burman); (iii) memoria puramente intelectual.[13] En su carta a Huygens, Descartes recurre a (iii) para aproximarse con la mera razón natural a un misterio de fe religioso, como el recuerdo de los seres queridos *post mortem*, y para intentar consolar a su amigo.

Conclusiones

Tanto en la conversación con Burman como en la correspondencia con Arnauld, Descartes expone su "doctrina" de la memoria intelectual, aunque parcialmente y con relativo detalle. Sin embargo, los argumentos de Descartes muestran claramente que los motivos que lo llevaron a postular una memoria que no reside en ningún órgano corpóreo son estrictamente filosóficos. En efecto, por un lado, la memoria corporal solo puede retener las impresiones sensibles causadas por las palabras, es decir, los sonidos de sus sílabas o las figuras de sus letras. Por consiguiente, lo único que puede proporcionar la memoria corporal es el recuerdo de los rasgos físico-sensibles de las palabras, pero no su significado. Por otro lado, las impresiones sensibles solo constituyen causas ocasionales que afectan a la mente de diversas maneras. Sin embargo, ninguna impresión sensible, por sí misma, es capaz de "comunicar" a la mente su novedad. Para detectar esto, se requiere un examen del intelecto, mediante el cual este último "compara" impresiones sensibles presentes y pasadas. En su entrevista con Burman, así como en sus cartas a Arnauld, Descartes analiza casos en los cuales

---

12　En su correspondencia, Descartes compara los vestigios fisiológicos de la memoria con los dobleces de un trozo de papel o de una prenda de vestir (AT IV, 114-15).

13　En su importante libro sobre Descartes, Martial Gueroult (1968) señala que es posible distinguir cuatro tipos de memoria en los escritos de Descartes: "memoria intelectual y memoria sensible, que no residen en el cuerpo; memoria corporal inscrita en el cerebro, memoria corporal esparcida en nuestros órganos y en nuestros músculos" (p. 41). Desafortunadamente, esto es todo lo que dice Gueroult sobre los tipos de memoria en Descartes, y no ofrece ningún argumento para defender esta cuádruple distinción. Lo que resulta llamativo de estas declaraciones, entre otras cosas, es que Gueroult reconoció una memoria sensible en Descartes que, sin embargo, no reside en el cuerpo del ser humano que la posee. Por consiguiente, esa memoria sensible solo puede residir en la mente.

memoria corporal e intelectual operan en conjunto para que el ser humano pueda acceder a diversos recuerdos.

Las razones que motivaron a Descartes parecen ser epistémicas, semánticas y relativas a su doctrina de la mente, no teológicas. Descartes no ha postulado una memoria intelectual simplemente para continuar con una tradición ortodoxa escatológica; no es plausible pensar que Descartes, al menos en este caso, se haya sentido forzado por una tradición teológica preexistente a postular una memoria intelectual.

Según la perspectiva de Descartes, los fieles disponen de dos vías complementarias para conocer la distinción de dos memorias de naturaleza completamente diferente: la gracia divina y la racionalidad natural. Sin embargo, a pesar de que la luz natural de la razón y la luz de la fe se complementan, no debe olvidarse la finalidad que tiene la revelación de Dios. En una carta de agosto de 1638, dirigida a un autor desconocido, Descartes afirma que "nuestros razonamientos no nos persuaden de ninguna cosa que sea contraria a lo que Dios ha querido que creyésemos" (AT II, 348). Ahora bien, Descartes añade que sería un error "querer sacar [de las escrituras santas] el conocimiento de verdades que solo pertenecen a las ciencias humanas, y que no sirven en absoluto para nuestra salvación" (AT II, 348). Tratar de extraer "verdades adquiridas" (AT II, 347) de las sagradas escrituras "es aplicar la escritura santa a un fin que Dios no le ha dado en absoluto" (AT II, 348). Como podemos obervar en esta carta, la finalidad de las escrituras santas es, a juicio de Descartes, nuestra salvación. De hecho, para Descartes, esta es la tarea principal de la teología. Como sostiene Descartes en el *Discurso del método*, la teología nos enseña fundamentalmente a "ganar el cielo" (AT VI, 6). Ahora bien, las enseñanzas de la filosofía (o ciencia por razón natural) y las enseñanzas de la teología no son contradictorias, sino complementarias. La teología extiende el conocimiento que la razón, por sí sola, puede alcanzar.[14] Por consiguiente, las motivaciones filosóficas que llevan a postular una memoria intelectual no constituyen un obstáculo para las motivaciones teológicas que podrían existir para postular dicha memoria. Las motivaciones filosóficas, aunque sean diferentes a las motivaciones teológicas, confirman la certeza teológica de la existencia de la memoria intelectual.

Referencias

Ariew, R. (2014). *Descartes and the first cartesians.* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199563517.001.0001

Ariew, R., Des Chene, D., Jesseph, M., Schmaltz, T. & Verbeek, T. (2015). *Historical dictionary of Descartes and cartesian philosophy.* Rowman & Littlefield.

---

14   Debido a esto, el conocimiento por fe religiosa no sería irracional, sino "supraracional".

Beyssade, J. M. (2001). *Études sur Descartes.* Seuil.

Coímbranses (1593). *Comentarii in libros Aristotelis qui parva naturalia appellantur.* s.e.

Cottingham, J. (1993). *A Descartes dictionary.* Blackwell.

De Buzon, F. & Kambouchner, D. (2005). *Le vocabulaire de Descartes.* Ellipses.

Descartes, R. (1897/1913). *Œuvres de Descartes* (11 volúmenes) . C. Adam & P. Tannery, (Eds.) Leopold Cerf.

Eustaquio de San Pablo. (1647). *Summa philosophica quadripartita de rebus dialecticis, moralibus, physicis, et metaphysicis.* s.e.

Gilson, E. (1979). *Index scolastico-cartésien.* J. Vrin.

Gueroult, M. (1968). *Descartes selon l'ordre des raisons,* Tomo 2. Auber.

Joyce, R. (1997). Cartesian memory. *Journal of the History of Philosophy, 35,* 375-393. https://doi.org/10.1353/hph.1997.0046

Sancti Thomae de Aquino (1889). *Opera omnia* (Tomus 5.). Iussu Leonis 13 P. M.

Smith, K. (2015). *The Descartes dictionary.* Bloomsbury.

Sutton, J. (2007). *Philosophy and memory traces. Descartes to connectionism.* Cambridge University Press.

Causation and mnemonic roles: on Fernández's Functionalism*

Nikola Andonovski

Université Grenoble Alpes, Grenoble, France
E-mail: ni.andonovski@gmail.com

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

Cómo citar este artículo

ARTÍCULO
DE REFLEXIÓN

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 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 I 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

The sense of mineness in personal memory: problems for the endorsement model*

Marina Trakas

Instituto de Investigaciones Filosóficas (IIF), SADAF/CONICET, Argentina
Email: marinatrakas@gmail.com

Abstract: What does it take for a subject to experience a personal memory as being her own? According to Fernández' (2019) model of endorsement, this particular phenomenal quality of our memories, their "sense of mineness", can be explained in terms of the experience of the mnemonic content as veridical. In this article, I criticize this model for two reasons: (a) the evidence that is used by Fernández to ground his theoretical proposal is dubious; and more importantly, (b) the endorsement model does not accommodate many non-pathological everyday memories that preserve their sense of mineness, but whose veridicality is explicitly denied, suspected, not automatically endorsed, or neither denied nor endorsed. Finally, I sketch two alternative explanations: one also problematic, the other one more promising, and present some normative advantages of the latter. This also displays the undesirability of the endorsement model from a normative perspective.

Keywords: personal memory, sense of mineness, sense of ownership, veridicality, non-believed memories

---

OPEN ACCESS

El sentido de mismidad de nuestros recuerdos personales: problemas del modelo de respaldo

Resúmen: ¿Qué significa que un sujeto siente un recuerdo personal como propio? Según el modelo de respaldo de Fernández (2019), esta cualidad fenomenal particular de nuestros recuerdos, su "sentido de mismidad", se puede explicar en términos de la experiencia del contenido mnemónico como verídico. En este artículo, critico este modelo por dos razones: (a) la evidencia que utiliza Fernández para fundamentar su propuesta teórica es dudosa y, aún más importante, (b) dicho modelo no acomoda muchos recuerdos cotidianos no patológicos que preservan su sentido de mismidad, pero cuya veracidad es explícitamente negada, sospechada, no respaldada automáticamente, o ni denegada ni avalada. Finalmente, esbozo dos explicaciones alternativas: una también problemática, la otra más prometedora, y presento algunas ventajas normativas de la última, lo que a su vez muestra la indeseabilidad del modelo de respaldo desde una perspectiva normativa.

Palabras clave: memoria personal, sentido de mismidad, sentido de posesión, veracidad, recuerdos no-creídos

Marina Trakas

PhD in Philosophy (Macquarie University, Australia); PhD in Cognitive Science (Institut Jean Nicod, France); Postdoctoral researcher, Consejo Nacional de Investigaciones Científicas y Tecnológicas (CONICET), Argentina. Member of the Instituto de Investigaciones Filosóficas (IIF), Sociedad Argentina de Análisis Filosófico / CONICET.

ORCID: 0000-0001-6977-140X

When we remember a past personal experience, we are not only aware of the fact that a certain type of mental phenomenon is being instantiated, in this case a personal memory, but we are also aware of the memory as being our own. These are in principle two different feelings. A hypothetical creature to whom we may transfer our personal memories could certainly have the sense that a memory is being presented to her, without experiencing that these are *her* memories, that what is presented to her are *her* past experiences. In our everyday mental life, the first kind of feeling is common to all mental phenomena, i.e., all non-pathological mental phenomena are presented as occurring in our minds; whereas the second feeling seems to be particular to personal memory experiences. We feel our personal memories as our own in a non-trivial essential way: they belong to us in a way in which they could not belong to someone other than us. They refer to our own past, to our previous experiences, and not to the past and experiences of someone else; that is why we attribute our memories to ourselves.

In this article, I focus on this particular sense in which personal memories are felt as our own, and analyse the specific conceptualization of this feeling recently developed by Fernández (2019). My aim here is relatively simple: I show that the "endorsement model" proposed by Fernández (2019) to account for this particular feeling does not offer a good explanation of the phenomenon it intends to explain. For this purpose, I briefly characterize the endorsement model and then explain in detail the problems that this model faces. I argue, first, that the endorsement model is not firmly grounded on the empirical evidence presented by Fernández, as he claims; and second, and more importantly, that it cannot account for the sense of mineness of a group of non-pathological everyday memories which are not endorsed by the rememberer. I use the term "sense of mineness" to refer to the explanandum of the model. Fernández is sometimes ambiguous about the terminology and indistinctly uses "sense of mineness" and "sense of ownership". I simplify the discussion by using a single term, without necessarily taking a stance about its suitability in comparison with other similar terms appearing in the literature.[1] After the analysis of Fernández' endorsement model of the sense of mineness of our personal memories, I introduce an alternative version of the endorsement model which nonetheless presents some similarities with Fernández' proposal. I show that this more modest version of the endorsement model is also problematic, and that a better explanation of the sense of mineness of our personal memories may emerge from an improved and more detailed version of the "identification model". I finally sketch some normative advantages of the latter option, which at the same time show the undesirability from a normative perspective of the endorsement model and its assumptions about the nature of memory.

---

1   In fact, in the current literature about mental states (and not, in particular, about memory), the terminology to refer to the general feeling of presence of a mental state in our consciousness is quite messy: sense of ownership, sense of autorship, (sense of) mineness, for-me-ness, felt mineness, etc. (see, for example, Gerrans, 2001; Salje & Geddes, forthcoming).

1 The endorsement model

The endorsement model of the sense of mineness is defended by Fernández (2019). According to his model, the sense of mineness of our memories, that is, the experience of a memory as being the subject's own, is the experience of the memory as matching the past. When a subject remembers, she feels that the remembered action or experience did take place in the past, so she endorses her memory and experiences it as being her own, that is, as having a sense of mineness: "the feeling of a memory as being one's own and the sense that the memory is matching the past are one and the same experience" (Fernández, 2019, p. 123).

This equation is possible because when the subject feels that a remembered scene has been experienced by her, she actually experiences one of the things represented by her memory: the fact that the memory has been caused by her having a perception (Fernández, 2019, p. 125). In fact, Fernández defends a reflexive view of memory: memory not only represents a past perceptual experience of the subject as having caused the memory, but it also represents that perceptual experience as having been veridical, so as a past perception, and not, let's say, as a past illusion or hallucination. It does not matter if the past perceptual experience was actually veridical or not; it matters how memory represents it, and according to Fernández memory represents it as being veridical. So memory represents its own causal history by representing itself as originating in a past perception. The experience of this specific causal origin is the experience of the sense of mineness. Note that the experience of the causality of the memory includes the experience of its veridicality: by representing its causal origin in a past perception and not simply in a past experience, a memory represents itself as being veridical, as matching the past. Fernández embeds the sense of mineness in the content of memory, and in this way grounds the phenomenology of memory in its intentionality, as he originally intended (Fernández, 2019, p. 23).

Therefore, within this framework, a memory that lacks a sense of mineness, a "disowned memory", is a memory whose content has been disrupted. The memory no longer represents itself as originating in a past perception; it represents itself as originating in a past perceptual experience, but this past perceptual experience is not represented as having been veridical. So the subject believes that she is remembering a scene, but she does not feel like she is remembering that scene, because she does not have the feeling that the scene represented in those memories happened in the past. Nonetheless, a disowned memory can still be recognized as a memory by the rememberer, because the content of disowned memories is sufficiently similar to the one of a typical memory. A disowned memory still represents part of its causal origin: that it has been originated in a past perceptual experience. So someone who disowns a memory may still refer to her mental state as memory and not as an

episode of imagination. The content of disowned memories is disrupted (Fernández, 2019, p. 125), but it is still "mnemonic" in nature because it is still reflexive.

2 Problems for the endorsement model

Fernández' proposal is attractive because of the coherence between the endorsement model of the sense of mineness and his theory of memory content. Despite this, it presents many shortcomings that make it an unsatisfactory account of the sense of mineness.

Fernández bases his whole theoretical analysis of the sense of mineness in the linguistic analysis of some reports of the patient R.B., investigated by Klein & Nichols (2012). R.B. is a subject who, after an accident, seemed to suffer a specific memory problem for some time: he claimed to have memories but at the same time to not feel these memories as his own. Fernández' model of sense of mineness as endorsement is presented as the best model that accommodates R.B.'s case. So it intends then to be an empirically grounded model that offers the best explanation of the sense of mineness of our memories. Nonetheless, Fernández' endorsement model is inadequate to explain the phenomenon it intends to explain, and this for many reasons.

First, there is no reason to believe that the endorsement model is the best model that accommodates R.B.'s case. Fernández bases his theoretical conclusions on the analysis of only *some* of R.B.'s reports cited in the literature (Fernández, 2019, p. 114). An account solely based on R.B. reports should provide an exhaustive and detailed analysis of all the available reports. But this is not Fernández' strategy, who deliberately omits those reports that he considers to be neutral for his analysis, but that may, in fact, not be neutral. Two reports omitted are of particular interest for his argumentation (Stanley & Klein, 2012, p. 685; p. 688):[2] both of them mention that R.B. did not feel as if he was the same person who performed or experienced something in the past. This kind of description of the phenomenon would give more support to the explanation given by Klein and Nichols (2012) and rejected by Fernández: the sense of mineness refers to the sense of numerical identity with the past person (which is known as "the identifi-cation model": Fernández, 2019, pp. 117-119). On the other hand, it would be difficult to interpret these descriptions according to the framework proposed by the

---

2   "I could answer any question about where I lived at different times in my life, who my friends were, where I went to school, activities I enjoyed, etc. But none of it was 'me'. It was the same sort of knowledge I might have about how my parents met or the history of the Civil War or something like that" (Klein & Nichols, 2012, p. 685). "I didn't feel 'down' about not being able to walk, etc. Because it was as if I was learning to walk for the first time. There was no sense of loss. Only a sense of gaining new skills and meeting these interesting new challenges. I knew that I once could walk, but it wasn't 'me' who once could walk" (Klein & Nichols, 2012, p. 688).

endorsement model. Fernández should have considered all R.B.'s verbal reports publicly available if he intended to use them in order to weigh one model against the other.

Second, it is doubtful that a good account of our sense of mineness is an account that accommodates R.B.'s case. It is not really a problem that R.B. is the single case worldwide reported with this specific pathology.[3] Many major contributions to cognitive science come from the study of single cases (Rosembaum, Gilboa & Moscovitch, 2014). The real problem is that there is not robust data on R.B. 's case, so it cannot be taken as a solid empirical ground from which to build a theory to explain memory phenomena. The loss of R.B.'s memories' sense of mineness was temporary (Klein & Nichols, 2012, p. 688), and he was only interviewed by these researchers. In fact, only a few fragments of the interviews are offered by the authors with the aim of defending more philosophical than empirical claims about the nature of personal identity, autonoesis, and episodic memory (Klein & Nichols, 2012; Klein, 2013; 2014; 2015), and at least one of these reports is not cited consistently in different publications.[4] Furthermore, Klein and Nichols mention verbal fluency and short-term memory tests performed on R.B. (Klein & Nichols, 2012, p. 684), but it does not seem that R.B. has been subject to exhaustive formal testing. R.B.'s case is blurred in so many ways that it is better to not count on it as solid evidence, or to take it with extreme caution.

Because of these two previous reasons, it is better to ignore the endorsement model's grounds on R.B.'s reports and to analyse it as a theoretical proposal that is not empirically grounded. This means that the equation between sense of mineness and memory endorsement needs to be tested with real life memories, in order to see if it is sufficiently explanatory or not. Two moves are possible for this purpose: a more analytical move and an empirically oriented move. As I show next, both of them finally meet and support the same conclusion, that is, that the endorsement model does not constitute a good account of the sense of mineness of our personal memories.

---

3   The two other historical cases invoked by Klein and Nichols that are supposed to be quite similar to R. B. are not really alike. Unlike R. B., both patients are amnesic. From the short description given by Talland (1964), his patient seems to present short-term memory deficits and source amnesia: after recalling some past personal events, he believed that the interviewer was the original source of the information that in fact he had given himself. J.V., the patient studied by Stuss and Guzman (1988), presents retrograde amnesia, especially for events personally experienced. He had nonetheless relearned some of his own past experiences through conversations and photographs, and he was aware in many cases of the real source of his "new memories". Because of this origin, these relearned events were recalled as if they were semantic information, and that is why "he has no feeling of personal warmth, intimacy, or belonging that the memory is his, or that he was somehow involved" (Stuss & Guzman, p. 27).

4   Compare "I can see the scene in my head. I'm studying with friends in the lounge at my residence hall. I am able to re-live it. I have a feeling [...] a sense of being at there, at MIT, in the lounge. But it doesn't feel like I own it. It's like I'm imagining, re-living the experience but it was described by someone else" (Klein & Nichols, 2012, p. 687), with "I can see the scene in my head [...] I'm studying with friends in the lounge in the residence hall. But it doesn't feel like its mine [...] that I own it. It's like imagining the experience, but it was described by someone else" (Klein, 2015, p. 18). In this second quote, there is no mention of "re-living".

The first move consists in finding a counterexample, and this is quite easy: we only have to look in our memory *repertoire*. Until I was more or less 20, I had two or three extremely vivid memories of my childhood that I still used to experience as being my own, as having a strong sense of mineness, but without endorsing their accuracy. Today they have lost their strong experiential character, as it has probably happened to all the memories from my childhood; but until my twenties these specific non-endorsed memories were extremely vivid and felt like real ones. One of them referred to a dinner at home with Enrique Discépolo, a famous tango composer. I used to remember him having dinner in our dinner room, with me and my parents. He was sitting in a particular chair near a corner, around an old dining table that we used to have when I was a young child, dining with my parents as if they were very close friends. When I was old enough, and after asking many times about this, my parents made me understand that this scene did not take place. Unfortunately, not only did my parents not personally know Discépolo, but Discépolo had died decades before my birth. So the event I remembered was not even remotely possible. Nonetheless, for a long time until my twenties, this past scene was so vivid and full of details, that I still experienced it as being my own, in a similar way I experienced the mineness of truly accurate memories. But because I knew this event was not real, I did not endorse the memory. I felt that I was remembering without feeling nor believing that the remembered experience did take place in the past and was veridical. But I still had the feeling that this scene belonged to *my personal* past in a non-trivial way and was *my* memory. This feeling could not possibly come from my experience of the fact that this memory has been caused by me having a perception. My memory did not anymore represent my past perceptual experience as having been veridical, so I could not possibly experience what was not represented in my memory, yet I still experienced a sense of mineness.

My Discépolo memory constitutes a good counterexample to Fernández' model of sense of mineness as endorsement. According to Fernández' framework, the content of my Discépolo memory would be disrupted: the memory does not represent itself as being caused by a past perception, so it does not represent itself as being veridical; but it still presents a sense of mineness. Therefore, the disruption or not of this particular memory content, and its consequent experience, does not seem to be related to the presence or absence of the sense of mineness.

My Discépolo memory provides evidence against the endorsement model proposed by Fernández. But this counterexample alone has no more value than R. B.'s reports, which Fernández considers to be supportive of his thesis. Nonetheless, it turns out that my Discépolo memory is not an isolated case, and that these cases are not rare phenomena. The psychologists Alan Scoboria and Giuliana Mazzoni have largely studied these kinds of memories of events that are no longer believed, which are called "non-believed memories". According to the first study on non-believed memories (Mazzoni, Scoboria & Harvey, 2010), more than 20% of participants reported a non-believed memory from a poll of more than 1500 students. Another more recent study, which

analysed "humorous, obscure and uncomfortable" false memories that have been spontaneously submitted by people for an art project, found that more than a half of those false memories were non-believed memories (Otgaar, Bücken, Bogaard,Wade, Hopwood, Scoboria & Howe, 2019). In fact, nonbelieved memories need not be false memories; true memories can also be non-believed, especially due to social influence (Mazzoni, Clark & Nash, 2014) So non-believed memories are quite common. Non-believed memories are autobiographical recollections that exist without accompanying autobiographical belief, or belief in the occurrence of the event remembered. Despite the withdrawal of belief, non-believed memories are still experienced as memories and not, for example, as imaginative acts, because they present very high ratings of memory-like qualities. Non-believed and believed memories show similar ratings of reliving, mental time travel, vividness of details, perceptual visual and tactile characteristics, intensity of feeling, richness of emotional content, and clarity of the location and spatial arrangements of objects and people in the event (Mazzoni et al, 2010). It is true that none of these characteristics correspond to the "sense of mineness" understood in a philosophical sense. But it is clear from all these characteristics that during recollection of non-believed memories, the rememberer is aware of the scene as having been experienced by *her* in the past. If she was not aware of the scene as having been experienced by her in the past, her current experience would not be taken to be a personal memory experience, so it would not present high ratings of reliving, mental time travel, vividness and clarity of details and elements. Although we may disagree about the necessity of all these traits for an experience to be a personal memory (some memories may not entail mentally travelling back in time, or may be poor in details), there is agreement that the presence of these traits signals a memory experience. Sufficient levels of vivid perceptual imagery, emotional content and a sense of re-experiencing the past generally lead the rememberer to label her mental representations as memories (Rubin, Schrauf & Greeberg, 2003). We may also disagree about the precise nature of the sense of mineness, but there is also agreement that the sense of mineness is a constitutive part of what a personal memory experience is. This point is nicely illustrated by Sacks (2005), who suggests that different phenomenal characteristics of our personal memories, including the sense of mineness, are deeply intertwined. In his retelling of the acknowledgement that a memory of his childhood was false, he states to not having found any phenomenal difference between a true and believed memory and his now false and nonbelieved memory: this latter, he tells us, "was equally clear, it seemed to me —very vivid, detailed, and concrete (...) it still seems to me as real, as intensely my own, as before" (p. 11). It is thus undeniable then that non-believed memories, like any kind of memory, present a sense of mineness. As Mazzoni, Scoboria, & Harvey (2010) explain, "there is no need to believe in the occurrence of an event to still 'feel' the memory as a compelling experience" (p. 12). Memories can be non-believed, that is, they can lack the veridicality postulated by the endorsement model, but still feel as

something that has been somehow previously experienced by me, as something so "strongly embedded in my psyche" (Sacks, 2005) and in my personal past that they cannot be phenomenally distinguished from believed memories.

It is also true that believed memories present greater personal significance and connectedness to other life events than non-believed memories, making non-believed memories less "connected to the self" than their counterpart (Mazzoni, et al., 2010, p. 13). But this does not undermine their sense of mineness. And if their sense of mineness is considered to be diminished because of this lack of connectedness, this diminished sense of mineness could not be taken as distinctive of non-believed memories. There are many believed memories which, for different reasons, could also be considered to be poorly connected to other past events or to the self. "Free radicals", that is, memories of events that have not become integrated with autobiographical knowledge (Conway, 2009), memories of specific past shameful acts remembered after a radical change in one's traits and values (Goldie, 2012), could both be considered to present a diminished sense of mineness understood in this sense. Nonetheless, there is no reason to consider the sense of mineness of a memory as equivalent to the experience of its connectedness and coherence with the self and other memories. Although this factor could influence the sense of mineness in a qualitative manner, it determines neither the absence nor the presence of the sense of mineness, as the phenomenology of some non-believed memories (such as my Discépolo memory, which was neither coherent nor possible) and of free radicals show. Absence of coherence with the self or other memories does not prevent the rememberer from feeling the past event recalled as having been experienced by her.

In consequence, it seems that Fernández' endorsement model of the sense of mineness does not pass the test: it is not sufficiently explanatory. The existence of non-believed memories leads us to the third and decisive point against it: the endorsement model does not accommodate some non-pathological everyday memories whose veridicality is explicitly denied but which still present a sense of mineness.

An advocate of the endorsement model proposed by Fernández could nevertheless raise an objection. Non-believed memories are still experienced as memories because in the past they actually represented themselves as originating in past perceptions and formerly held a high belief value. So the persistence of the sense of mineness can be explained in terms of the original content of the memory, which represented itself as being veridical. Although this content has changed, i.e., it does not represent anymore as originating in a past perception and being veridical, the sense of mineness of the memory still persists as a residual condition of the experience of this original content. According to this view, the endorsement model would account for the sense of mineness of non-believed memories; that is why it would be sufficiently explanatory for this phenomenon.

This line of argument is nevertheless problematic. It assumes that memories always intrinsically represent themselves as originating in past perceptions,

and so, that recollecting an event and believing that the event occurred happen simultaneously because they are either both part of the same experience, or at least strongly correlated. The functionalist theory of memory proposed by Fernández explicitly states this point: "For any subject S and proposition p, S remembers that p just in case S has some mental image i such that i tends to cause in S a disposition to believe both that p and that S experienced that p, and i tends to be caused in S by having experienced that p" (Fernández, 2019, p. 49). Tulving had also informally suggested the same idea: "the feeling that the present recollective experience refers to a past event, and the feeling that the experience is veridical, that it represents the past faithfully, are given as an integral part of the subjective experience of remembering" (Tulving, 1983, p. 187). But recent empirical research shows otherwise. Although prior research was inclined to analyse memories and autobiographical beliefs as partially dissociated but generally correlated (Scoboria, Mazzoni, Kirsch & Relya, 2004), new research using Confirmatory Factor Analysis[5] supports the idea that they are better conceived as strongly independent of each other (Scoboria, Jackson, Talarico, Hanczakowski, Wysman & Mazzoni, 2014). Memories and autobiographical beliefs, that is, beliefs in the occurrence of the event remembered, represent two distinct continuous variables that can manifest in different combinations across mental events. At the extreme of these possible combinations, there are four "pure" event types: "believed memories", "non-believed memories", "believed-not-remembered" and "not-believed-not-remembered". Whereas the first two refer to memory experiences and the third to autobiographical beliefs about personal events that are not remembered, the fourth refers to past personal events that may or may not have happened, but are neither believed nor remembered. This means that our everyday memories scatter across the two extremes: "non-believed memories" and "believed memories". Believed memories are certainly more common than non-believed memories, but between these two extremes there are many memories high in recollection for which the belief in the event remembered has been brought into question but not substantially undermined. These "grain-of-doubt" memories (Scoboria, Nash & Mazzoni, 2016)[6] occur frequently and, if the

---

5   Confirmatory Factor analysis (CFA) is an analysis tool that helps the researcher to test the fit of data to a previously specified model, and thus bridge the gap between observation and theory. As Scorobia & et al. (2014) explain, "the researcher defines in advance what factors will be present and onto which specific factor(s) items will load. The data are then used to evaluate the extent to which the model explains covariance amongst the measured variables, resulting in an estimate of the degree to which the proposed theoretical model represents a good fit to the data" (p. 1246). Whereas a model with different items loading on a single factor (memory = belief) did not fit the data well, a two factor model (memory and belief) was a good fit of the data. The same procedure was used to assess the data-model fit of the different event types mentioned below in the main text.

6   Scoboria, et al., (2016) have in fact called these memories "grain-of-doubt" non-believed memories when analyzing the different types of non-believed memories. I think nevertheless that it would be better suited with their previous work (Scoboria, et al., 2014) to talk about "grain-of-doubt" memories *tout court*; technically, these memories are not "non-believed", but "suspected" or "dubious".

need to know their truth value arises, this need is generally resolved by turning to social resources for assistance, such as external evidence and testimony. So "grain-of-doubt" memories are another good example of an explicit dissociation between memory and autobiographical belief.

Furthermore, this dissociation is also observed between the predictors of the variables (Scorobia & et al., 2014). Whereas rich visual imagery, the feeling of reexperiencing, emotion intensity, and event specificity are good predictors of recollection but do not predict belief, general event plausibility and personal plausibility predict autobiographical belief strongly and recollection only weakly. In fact, autobiographical beliefs seem to be influenced by multiple sources of information and processes, such as factual knowledge, inferential processes and, particularly, by the acquisition of socially transmitted information and social influence (Scoboria, Boucher & Mazzoni, 2015). Recollection is only one among many factors that influence autobiographical belief, and it is not even a necessary one. As Scorobia et al., (2014) have stated, "the processes that impact recollection tend to be internal and cognitive/memorial in nature", whereas "the processes that impact autobiographical belief originate in both internal and external experience and are strongly socially mediated" (p. 1255). Even the metacognitive processes that govern the endorsement of an event as occurred differ from the metacognitive processes that govern the endorsement of an event as remembered. The conclusion that one remembers an event depends on the interaction between the qualities of the "raw output" of memory retrieval processes (vividness, fluency, etc) and the implicit criterion for judging that mental content as a memory. On the other hand, the metacognitive processes that determine the likelihood that a personal event happened can be based on information that the rememberer already has —which can but need not be the memory of the event in question— and on newly provided information (Mazzoni & Kirsch, 2002). Certain metacognitive feelings, such as the feeling of rightness, confidence, error or uncertainty, play a crucial role in motivating (or not) the subject to resort to social sources (Arango-Muñoz, 2013). When people turn to social sources, they can (a) endorse the new information; (b) reject the new information; (c) seek additional information; (d) diminish the importance of the event, especially when motivations to maintain close and meaningful relationships override motivations to be accurate (Scorobia et al., 2014). While in cases (a) and (b) an autobiographical belief is formed or changed, in case (c) the subject remains in a state of doubt and inquiry, and in case (d) the belief or not in the occurrence of the event is withheld or suspended (the subject neither believes nor disbelieves the occurrence of the event).[7] In all four cases, nonetheless, a memory of the event in question may be present and persist.

---

7   For the notion of "suspended judgement" as a genuine third option between belief and disbelief, see, for example, Friedman (2013): "Suspending then is (other things equal) a way of (at least temporarily) terminating a deliberative process that is sufficient for getting into a state of agnosticism" (p. 179).

Therefore, non-believed memories, mainly of episodes of our childhood, are not the only case where the dissociation and independence between memory and belief is undeniable. Grain-of-doubt memories and "suspended-belief" memories are very common phenomena that happen in everyday life, for example, when we vividly visualize ourselves locking the front door of our house, but we are not really sure if we did it, so we go back to check that the door is locked indeed. Or when we clearly remember telling our partner to buy bananas, but she claims we did not, and in order to avoid a fight we simply say "ok, it does not matter", implicitly suspending our belief or disbelief in our past request even though it is still remembered. What is more, Blank (2016) mentions other memory phenomena where the divergence between memory and belief is also evident: the déjà vu phenomenon (feeling of recollection + a strong awareness of inaccuracy); some cases of misinformation effect (feeling of recollection of an event + belief in a different event due to the information presented later); some memories of childhood sexual abuse (no memory of abuse + belief in abuse; or memory abuse + disbelief in abuse); stereotype-induced memory distortion (vague and ambiguous memories + belief formed due to stereotypes). The independence of these two constructs, memory and belief, does not mean that they do not influence each other, although more studies are needed in order to better specify the nature of this interaction.[8]

In conclusion, the possible objection of the advocate of the endorsement model previously mentioned does not work. The sense of mineness of non-believed memories cannot be explained in terms of the persistence of the experience of a supposed original —but lost—mnemic content that represented itself as originating in a past perception and being veridical. Recent empirical evidence points to a different direction: the sense of mineness, which is intrinsically tied to the recollective experience, is a different construct from the degree of belief attached to a past remembered experience. Experiences remembered do not necessarily represent themselves as having been veridical in the past, as having been a past perception. At most, they represent themselves as having been somehow experienced *me*, the rememberer. But memories do not come with an intrinsic label that says "veridical experience" and that makes us feel that they were veridical; this label is ascribed to the past experience through different processes and evidence that are not necessarily mnemonic in nature.

---

8   Mazzoni & Kirsch (2002) suggest that, besides the repetition of the content, a belief may also create a false memory by affecting the person's criterion for judging the mental content to be a memory: "the more likely the event is to have happened, the easier it is to conclude that a corresponding mental content is a memory, rather than merely a fantasy, imagining, etc." (p. 140). Recently, it has been suggested that false memories can be reversed through different techniques that ultimately change the belief of the subject (Oeberst, Wachendörfer, Imhoff & Blank, 2021). Nonetheless, Hyman (2021) is sceptical about this prospect. According to him, what Oeberst et al. have shown is that the belief in the memory has been reversed but not the memory itself: "The memory remains. We may no longer believe the memory is true. But the memory remains. We can erase the belief in the memory. But the images and the stories may stay in memory" (para. 16).

The third and decisive point against Fernández' endorsement model can thus be reformulated in the following terms: the endorsement model is not sufficiently explanatory because it does not accommodate many non-pathological everyday memories that preserve their sense of mineness, but whose veridicality is explicitly denied, suspected, not automatically endorsed, or neither denied nor endorsed. Furthermore, the previous analysis shows another weakness of Fernández' proposal: the idea that the memory content represents itself as being veridical. The veridicality of a memory is a property that the memory may acquire when the event remembered is judged to have occurred. It is thus not internal to the content of the memory, but attributed after memory retrieval, through processes that are not necessarily mnemonic. So this not only rules out epistemic theories of memory, which equate memory to knowledge or belief, but also challenges Fernández' functionalism: although it is certainly true that many memories tend to cause in us a disposition to believe in the occurrence of the event remembered, the memory status of a mental state cannot be defined in terms of this disposition. Some memories do not produce a tendency to believe in their content.

3 Future research directions

All these previous considerations show that the endorsement model of the sense of mineness proposed by Fernández (2019) is not a good model of the phenomenon it intends to explain:

(a) There is no reason to believe that it is the best model of the sense of mineness that accommodates R. B.'s case;

(b) It is doubtful that a good account of our sense of mineness is an account that accommodates R. B.'s case;

(c) It does not accommodate many non-pathological everyday memories that preserve their sense of mineness, but whose veridicality is explicitly denied, suspected, not automatically endorsed, or neither denied nor endorsed.

The unsatisfactory explanatory power attributable to the endorsement model of the sense of mineness opens up different possibilities. On one hand, the sense of mineness of memories could still be conceived as the experience of the memory content; but this would entail a redefinition of the memory content. The memory content as it has been conceptualized by Fernández cannot explain the sense of mineness —and additionally, it does not seem to be a good conceptualization of the memory content. This move may nevertheless be tricky: it may demand the memory content to do "too much" explanatory work, especially when considering that memory is sometimes used for non mnemonic purposes, such as imaginative purposes. But this move still may be worth trying if the memory content is not considered as the simple activation

of memory traces (which can be used for non mnemonic purposes), but as the final output resulting from retrieval memory processes. Therefore, a more modest model of endorsement could be proposed to explain the sense of mineness. In this version of the model of endorsement, the experience of the sense of mineness still corresponds to the experience of the causal origin of our memories that is given in the memory content. But the memory content represents itself as originating in a past perceptual experience, not in a past perception, so it does not represent itself as being veridical. The memory content is still reflexive, that is, it represents its own causal origin, but it is neutral about the veridicality of the past experience that caused it. This more modest model of endorsement that is grounded on a more modest model of the reflexive nature of memory could still accommodate the sense of mineness. The sense of mineness would correspond then to the awareness of the scene remembered as having been experienced in the past. So in this moderate model of the sense of mineness as endorsement, what is endorsed is not the veridicality of the memory content, but simply the existence of a previous experience (perception, illusion, hallucination, dream, imagination, etc.) where the memory originated.

Nonetheless, this moderate model of endorsement may not be a good explanation of the sense of mineness. Consider again my Discépolo memory. I probably discovered Discépolo through a book in our library, which presented hand-made portraits of different tango composers together with their biography and the lyrics of one of their songs. Maybe then I dreamt about a dinner with Discépolo at home, but I am not really sure if my memory was due to a source monitoring error. Maybe there was no dream, no previous hallucination or other experience from which my memory originated, and my memory was the result of anomalous processing of different elements: my memory of the hand-made portrait of Discépolo, my parents listening to tango, my eagerness (?) to receive a celebrity at home, etc. The truth is that I do not have any idea about the kind of experience, if there was such experience, that engendered my Discépolo memory. So, although not technically impossible, it is suspicious to ground the sense of mineness in the awareness of the scene remembered as having been caused by a previous indeterminate, unspecified —and maybe non-existent— experience in the past.

In Fernández' model of endorsement as well as in this more modest proposal, the weight of the causal origin is put on the side of the object, i.e., the past event or experience, and not on the side of the subject. Although the experience is not always there, like in my Discépolo memory, the subject always is, and focusing on this aspect of the memory may be a more promising line of inquiry to explain the nature of the sense of mineness. Why do I feel particular mental states as *my* memories? Because they (re)present themselves as originating in my past self, who is the same as me. The idea that the sense of mineness refers to the sense of numerical personal identity with the past person was the original thesis proposed by Klein & Nichols (2012) to explain R. B.'s case, which was nevertheless rejected by Fernández for not being a good explanation of R. B.'s reports. It deserves nonetheless to be better explored, considering that there

are some unclear aspects that need further elaboration. To name just a few: it is not evident that the notion of identification with the past self properly captures the nature of the sense of mineness. The notion of identification implies a two-step process, or an inferential process that does not seem to be part of the experience of the sense of mineness, which is given immediately and non-inferentially.[9] But abandoning the notion of identification leaves an explanatory void that needs to be fulfilled with some other concept or explanation. On the other hand, another aspect that needs clarification refers to the "mental location" of the past self. Either the past self is —implicitly or explicitly— represented in the content of the memory, so the memory represents itself as originating in the past self and the sense of mineness arises from the experience of this content; or the past self is not represented in the content. In the latter case, the memory presents itself as originating in my past self, but this causal origin and the identification or "sameness" between present self and past self are not part of the memory content. They may be the phenomenal flip side of subpersonal and automatic metacognitive memory processes. In this framework, the sense of mineness of memories could be then conceived as a metacognitive feeling. This is probably a more viable option than assuming that the sense of mineness originates in the experience of the memory content itself, and a promising line of research that deserves further development.

4 Desirability of the endorsement model?

To conclude, I would like to briefly present some advantages of the ideas sketched in the previous section in order to explain the nature of the sense of mineness of our personal memories. These advantages show, at the same time, that Fernández' endorsement model of the sense of mineness as well as its assumptions about the nature of the memory content are not a desirable account of memory from a normative perspective.

First, when memory is conceived as different from belief or from a tendency to produce a belief, the rememberer understands that some beliefs she may have about past events personally experienced may have been formed from sources different from memory itself. This kind of understanding turns out to be useful in many settings in which what matters is the original memory, such as in forensic contexts. For example, a subject may believe that she saw a person X stealing a car, but in fact this may be a belief she formed from many sources different from the original memory: her previous knowledge of X's reputation as a thief and X usually wearing a cap; a neighbor telling her that he saw X that morning walking around the neighborhood in a suspicious way; a neighbor asserting that it must be X who stole

---

9   A similar criticism has been raised by McCarroll (2018) in his account of the way one sees oneself in visual observer memories.

Estud.filos n.º 64. Julio-diciembre de 2021 | pp. 155-172 | Universidad de Antioquia | ISSN 0121-3628 | ISSN-e 2256-358X   169

https://doi.org/10.17533/udea.ef.n64a08

the car. The only thing that she may actually remember is that she saw a silhouette of a man with a cap stealing the car.

Second, when memory is understood as not being necessarily veridical, the rememberer can pay more attention to her metacognitive feelings (especially feelings of uncertainty and error) and can be more motivated to resort to external and social sources if her memory is considered to be doubtful. A healthy dose of doubt triggers further inquiry and improves the chances of forming a better representation of the past. These kinds of epistemic actions can also be triggered by the understanding that in many circumstances memory does not present to us objective facts but subjective personal perspectives of what happened, and that other perspectives may be needed to better grasp the past. This kind of understanding may be particularly useful when remembering experiences that are important for interpersonal relationships. For example, if my partner remembers that I have been treating him with disdain but I do not, I may consider that *my* memory only presents to me *my* subjective perspective of the past, and that my partner's perspective may be not only a different perspective but also a valid perspective that needs my full consideration in order to get a complete picture of my past interactions with him (Trakas, 2019). Therefore, a sense of mineness understood as focused on the side of the subject ("it was me who experienced some event") may be more epistemically beneficial than a sense of mineness understood as focused on the side of the object ("this event happened" or "this experience took place"). The understanding of the subjective side of memories may help the subject to avoid or overcome self-deception, and to adopt certain kinds of attitudes such as being open-minded and responsive to others' memories, all of which is essential to become a more epistemically responsible rememberer.

References:

Arango-Muñoz, S. (2013). Scaffolded memory and metacognitive feelings. *Review of Philosophy and Psychology, 4*(1), 135-152. https://doi.org/10.1007/s13164-012-0124-1

Conway, M. A. (2009). Episodic memories. *Neuropsychologia, 47*(11), 2305-2313. https://doi.org/10.1016/j.neuropsychologia.2009.02.003

Fernández, J. (2019). *Memory: a self-referential account.* Oxford University Press. https://doi.org/10.1093/oso/9780190073008.001.0001

Friedman, J. (2013). Suspended judgment. *Philosophical studies, 162*(2), 165-181. https://doi.org/10.1007/s11098-011-9753-y

Hyman, I. (2021, May). Can we erase false memories? *Psychology Today.* https://www.psychologytoday.com/us/blog/mental-mishaps/202105/can-we-erase-false-memories

Gerrans, P. (2001). Authorship and ownership of thoughts. *Philosophy, Psychiatry, & Psychology, 8*(2), 231-237. https://doi.org/10.1353/ppp.2001.0009

Goldie, P. (2012). *The mess inside: narrative, emotion and the mind.* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199230730.001.0001

Klein, S. B. & Nichols, S. (2012). Memory and the sense of personal identity. *Mind, 121,* 677-702. https://doi.org/10.1093/mind/fzs080

Klein, S. B. (2013). Making the case that episodic recollection is attributable to operations occurring at retrieval rather than to content stored in a dedicated subsystem of long-term memory. *Frontiers in Behavioral Neuroscience, 7,* 3. https://doi.org/10.3389/fnbeh.2013.00003

Klein, S. B. (2014). Autonoesis and belief in a personal past: an evolutionary theory of episodic memory indices. *Review of Philosophy and Psychology, 5*(3), 427-447. https://doi.org/10.1007/s13164-014-0181-8

Klein, S. B. (2015). What memory is. *Wiley Interdisciplinary Reviews: Cognitive Science, 6*(1), 1-38. https://doi.org/10.1002/wcs.1333

Mazzoni, G. A. L., Clark, A. & Nash, R. A. (2014). Disowned recollections: denying true experiences undermines belief in occurrence but not judgments of remembering. *Acta Psychologica, 145,* 139-146. https://doi.org/10.1016/j.actpsy.2013.11.007

Mazzoni, G. A. L., Scoboria, A. & Harvey, L. (2010). Non-believed memories. *Psychological Science, 21,* 1334-1340. https://doi.org/10.1177/0956797610379865

McCarroll (2018). *Remembering from the outside.* Oxford University Press. https://doi.org/10.1093/oso/9780190674267.003.0001

Mueller, R. O. & Hancock, G. R. (2001). Factor analysis and latent structure, confirmatory. In N. J. Smelser & P. B. Baltes (Eds.), *International Encyclopedia of the Social & Behavioral Sciences* (pp. 5239-5244). Elsevier. https://doi.org/10.1016/B0-08-043076-7/00426-5

Oeberst, A., Wachendörfer, M. M., Imhoff, R. & Blank, H. (2021). Rich false memories of autobiographical events can be reversed. *Proceedings of the National Academy of Sciences, 118*(13), e2026447118. https://doi.org/10.1073/pnas.2026447118

Otgaar, H., Scoboria, A. & Mazzoni, G. A. L. (2014). On the existence and implications of nonbelieved memories. *Current Directions in Psychological Science, 23*(5), 349-354. https://doi.org/10.1177/0963721414542102

Otgaar, H., Bucken, C., Bogaard, G., Wade, K. A., Hopwood, A. R., Scoboria, A. & Howe, M. L. (2019). Nonbelieved memories in the falsememory archive. *Journal of Applied Research in Memory and Cognition, 8*(4), 429-438.

Rosenbaum, R. S., Köhler, S., Schacter, D. L., Moscovitch, M., Westmacott, R., Black, S. E. & Tulving, E. (2005). The case of KC: contributions of a memory-impaired person to memory theory. *Neuropsychologia, 43*(7), 989-1021. https://doi.org/10.1016/j.neuropsychologia.2004.10.007

Rubin, D. C, Schrauf, R. W. & Greenberg, D. L. (2003). Belief and recollection of autobiographical memories. *Memory & Cognition, 31,* 887-901. https://doi.org/10.3758/BF03196443

Estud.filos n.º 64. Julio-diciembre de 2021 | pp. 155-172 | Universidad de Antioquia | ISSN 0121-3628 | ISSN-e 2256-358X

171

https://doi.org/10.17533/udea.ef.n64a08

Sacks, O. (2005). On memory. *The Threepenny Review*. https://www.threepennyreview.com/samples/sacks_w05.html

Salje, L. & Geddes, A. (forthcoming). Conscious experience: what's in it for me? In M. García Carpintero & M. Guillot (Eds.), *The Sense of Mineness*. Oxford University Press.

Scoboria, A., Boucher, C. & Mazzoni, G. A. L. (2015). Reasons for withdrawing belief in vivid autobiographical memories. *Memory, 23,* 545-562. https://doi.org/10.1080/09658211.2014.910530

Scoboria, A., Jackson, D. L., Talarico, J., Hanczakowski, M., Wysman, L. & Mazzoni, G. A. L. (2014). The role of belief in occurrence within autobiographical memory. *Journal of Experimental Psychology: General, 143*(3), 1242–1258. https://doi.org/10.1037/a0034110

Scoboria, A., Mazzoni, G. A. L., Kirsch, I. & Relya, M. (2004). Plausibility and belief in autobiographical memory. *Applied Cognitive Psychology, 18,* 791-807. https://doi.org/10.1002/acp.1062

Scoboria, A., Nash, R. A. & Mazzoni, G. A. L. (2016). Sub-types of nonbelieved memories reveal differential outcomes of challenges to memories. *Memory, 25*(7), 876-889. https://doi.org/10.1080/09658211.2016.1203437

Stuss, D. T. & Guzman, D. A. (1988). Severe remote memory loss with minimal anterograde amnesia: a clinical note. *Brain and Cognition, 8*(1), 21-30. https://doi.org/10.1016/0278-2626(88)90036-X

Talland, G. A. (1964). Self-reference: a neglected component in remembering. American Psychologist, *19*(5), 351-353. https://doi.org/10.1037/h0039768

Trakas, M. (2019). On epistemic responsibility while remembering the past: the case of individual and historical memories. *Les ateliers de l'éthique / The Ethics Forum, 14*(2), 240-273. https://doi.org/10.7202/1071139ar

Tulving, E. (1983). *Elements of Episodic Memory*. Oxford University Press.

SIMPOSIO

Feeling the past: beyond causal content*

Gerardo Viera

University of Sheffield, Sheffield, United Kingdom
E-mail: gerardo.viera@gmail.com

Abstract: Memories often come with a feeling of pastness. The events we remember strike us as having occurred in our past. What accounts for this feeling of pastness? In his recent book, *Memory: A self-referential account,* Jordi Fernández argues that the feeling of pastness cannot be grounded in an explicit representation of the pastness of the remembered event. Instead, he argues that the feeling of pastness is grounded in the self-referential causal content of memory. In this paper, I argue that this account falls short. The representation of causal origin does not by itself ground a feeling of pastness. Instead, I argue that we can salvage the temporal localization account of the feeling of pastness by describing a form of egocentric temporal representation that avoids Fernández's criticisms.

Keywords: memory, temporal experience, perception, causation, time

---

Cómo citar este artículo

OPEN ACCESS

Sintiendo el pasado: más allá del contenido causal

Resumen: Los recuerdos a menudo vienen acompañados de un sentimiento de pasado. Los eventos que recordamos nos parecen haber ocurrido en nuestro pasado. ¿Qué explica este sentimiento de pasado? En su libro reciente *Memory: A self-referential account,* Jordi Fernández sostiene que el sentimiento de pasado no puede fundamentarse en una representación explícita del pasado del hecho recordado. En cambio, sostiene que el sentimiento de pasado se basa en el contenido causal autorreferencial de la memoria. En este artículo, sostengo que esta explicación no es satisfactoria. La representación del origen causal no fundamenta por sí misma un sentimiento de pasado. En cambio, sostengo que podemos salvar la explicación de la localización temporal del sentimiento de pasado al describir una forma de representación temporal egocéntrica que evita las críticas de Fernández.

Palabras clave: memoria, experiencia temporal, percepción, causalidad, tiempo

Gerardo Viera

Is a philosopher of cognitive science, who arrived at the University of Sheffield in the Autumn of 2020. Before coming to Sheffield, he was a postdoctoral researcher at the Centre for Philosophical Psychology at the University of Antwerp. Prior to this, Gerardo received his PhD in philosophy from the University of British Columbia.

ORCID: 0000-0002-3183-2294

1. Introduction

Pick a memory. For me, perhaps primed by one of Fernández's examples, it's a memory of struggling against the waves when I was a kid. I remember flailing around trying to stay afloat. I remember inhaling water and being scraped against the rocks as the waves hit. There was a mild sense of panic and I kept darting my eyes towards the beach. Now, imagine a qualitatively similar situation. Not one from your past but one that's completely imaginary. The flailing, the inhaling of water, the sense of panic are all the same between these events that are brought to mind. Yet, there seems to be a phenomenal difference between our experience of remembering and our experience of imagining. One way of characterizing this difference is in terms of our temporal phenomenology. Remembering involves a feeling of pastness directed at the remembered event that is missing in imagination. Memory seems to mentally transport us to our personal past. Imagining lacks this feature.

What accounts for this feeling of pastness? In his recent book *Memory: a self-referential account* (2019), Jordi Fernández develops a comprehensive account of mnemonic content that is supposed to answer this question. Memories, according to Fernández, are self-referential. They represent their own causal origins. According to Fernández, the feeling of pastness that is often associated with episodic memory is grounded in this self-referential causal content rather than any explicit representation of time.[1] In this paper, I will argue that this account falls short and fails to explain how the feeling of pastness can be grounded in the representation of causal origin.

The paper goes as follows: Section 2 describes the self-referential theory of episodic memory and how it is supposed to account for the feeling of pastness. This section also lays out Fernández's arguments against temporal localization views of the feeling of pastness in which the feeling of pastness arises from an explicit representation of the temporal location of remembered events. Section 3 argues that the self-referential explanation of the feeling of pastness falls short. It fails to explain why a representation of cause would give rise to a feeling of pastness without introducing an explicit representation of time. However, if Fernández's arguments against temporal localization views are correct, then there is no representation of time that could account for the feeling of pastness. Therefore, it would appear that we would have no options available to explain the feeling of pastness. In section 4, I argue that we can avoid this conclusion by appealing to a form of egocentric temporal representation, what I call *path-dependent representation*, that Fernández does not consider.

---

[1]   Fernández takes the feeling of pastness to be an essential component of episodic memories (or more accurately, episodic remembering). In this paper, I will not commit to that strong claim. Depending on what mental states we want to count as episodic memories we may take this feeling of pastness to be essential or not (see Boyle, 2020).

2. The self-referential account

Fernández's book covers an impressive range of topics in the philosophy of memory. The cornerstone of the book is his *self-referential account* of mnemonic content. I cannot do justice to the many arguments for this position raised in the book. Instead, I will simply present the view and see whether it can do the work he wants it to. Here is the self-referential account of mnemonic content:

> For any subject S, memory M and proposition q: If S has M and S would express M by saying that they remember that q, then there is a perceptual experience P that S would express by saying that they perceive that q, such that the content of M is the proposition {W: in W, M is caused by S having perceived that q through P} (Fernández, 2019, p. 78).

Memories, according to the self-referential account, represent themselves as being caused by one of the subject's perceptual experiences. My memory of struggling against the waves represents itself as having been caused by a perceptual experience of that struggle. The memory will be accurate just in case I really did have a perceptual experience of that struggle and my current memory is a causal effect of that perceptual experience. Contrast that with my imagination of a similar event. The act of imagining that event does not represent itself as being caused by any perceptual episode (even if my capacity to imagine this sort of scenario is partly caused by my having certain experiences).[2]

How then does this help us account for the feeling of pastness? Let's begin by getting clear on this phenomenological explanandum. Here is Fernández's characterization of the feeling of pastness:

> (PAST) For any subject S and proposition p: If S remembers that p, then S is aware of the fact that p as obtaining in the past (Fernández, 2019, p. 87).

It's this awareness that is our phenomenological target. Since explanations of phenomenological features of experience are notoriously problematic, its useful to specify an assumption in the book. Fernández adopts a form of representationalism according to which the phenomenal features of episodic remembering *depend* on the representational features of episodic memory. The relation is a fairly weak one. As he puts it, "the content of a memory will *give us some information* about the way in which the memory feels to the subject" (Fernández, 2019, p. 29; emphasis added). For the purposes of this paper, I will simply accept this form of representationalism.

---

2   Fernández allows that a memory can fail to accurately represent its causal origin. What matters for something's being a memory is the represented causal relation, not the actual causal relation.

Before seeing how the self-referential view accounts for PAST, it's useful to see why Fernández thinks that we cannot account for PAST by appealing to what I call *temporal localization accounts* in which episodic memories represent the temporal location of remembered events. Fernández considers two forms of temporal localization —self-independent and self-dependent accounts. Let's take these in turn.

SELF-INDEPENDENT TEMPORAL LOCALIZATION. According to this approach, memories locate events in time by placing events at subject-independent temporal locations. My memory of struggling to swim might represent that event *as occurring on 15/07/1991.* Or perhaps it represents the event *as occurring in the early 90s.* No reference to the present is made. According to Fernández, this form of temporal localization cannot account for PAST. Consider a case where I imagine what will happen in the future. I might imagine myself struggling to swim and imagine this event *as occurring on 15/07/2029.* Given that it is currently 2021, the imagined event is in the future and the remembered event is in the past. However, nothing in the content of these mental states captures the pastness of the remembered event or the futureness of the imagined event. Some other content is needed that takes into account my current temporal perspective.

SELF-DEPENDENT TEMPORAL LOCALIZATION. The self-dependent account builds in a temporal perspective into the content that locates remembered events in time. Now, in 2021, when I remember my struggling to swim in 1991, that memory may represent that event *as having occurred 30 years ago.* Or perhaps it represents the event *as having occurred 30-ish years ago,* or it just represents the event *as having occurred some distance in the past.* In all these cases, the memory represents the event as being a certain temporal distance in the past relative to my current temporal perspective. This sort of content would account for PAST, but Fernández argues that this account violates certain intuitions about what counts as an accurate memory.

Consider the following scenario. I struggled to swim in 1991. It's now 2021 and I'm given the opportunity to travel back in time to 1989. When I travel back I remember that moment where I struggled to swim. It still has its distinctive feeling of pastness. However, according to the self-dependent account of temporal localization, this feeling of pastness is the result of that memory locating the event as being in the past. Is this experience accurate? It seems that a time traveler should still be able to successfully remember their childhood experiences. However, Fernández argues this isn't possible given the self-dependent form of temporal localization. My memory represents the event as being in the past, but given my time travelling to 1989, the event is not actually in the past but rather it is in the future. The account takes intuitively accurate memories and renders them inaccurate. This is a problem.[3]

---

3   Fernández is aware that time travel leads to all sorts of counterintuitive conclusions. For this paper, we can set aside these concerns.

Fernández concludes that we can't account for PAST through temporal localization. The self-referential account, however, supposedly avoids these worries. Memories represent their causal origin, not the temporal location of the remembered event or perceptual experience. Since in the actual world causes precede their effects, the representation of a memory as being the causal effect of a perceptual experience "goes hand in hand with (that remembered event/perceptual experience) having a certain position in time, namely, being in the past" (Fernández, 2019, p. 108).

It's at this point that there is a certain ambiguity in Fernández's account. If memories represent their causal origins, and not the temporal location of remembered events, then were we simply mistaken in describing the distinctive phenomenology of episodic memory as a *feeling of pastness*? Wouldn't the resulting phenomenology be one of causal origin since that is what is represented? It would appear that Fernández is giving an error theory of PAST. We actually experience a feeling of causation, but because of the close connection between time and causation, we misdescribe this experience of causation as an experience of pastness. Yet, this isn't how he describes his position. He takes the self-referential account of memory to ground an awareness of time.

In a later paper, Fernández (2020) explicitly rejects this error theory in a response to a criticism by Denis Perrin (2018). Perrin argues that if Fernández is correct, that memories represent their causal origin and not the pastness of the remembered event, then the pastness of the remembered event must be inferred or derived from this content, and therefore the feeling of pastness is derived from the feeling of causation. But, Perrin continues, this can't be the case since the feeling of pastness is an intrinsic feature of episodic remembering. The aim here isn't to assess Perrin's argument. Rather, it's Fernández's response that informs us about his rejection of the error theory and how he understands the relationship between the representation of causal origin and PAST. Fernández (2020), agrees with Perrin that the feeling of pastness is not inferred from the content of memory. Instead, he argues that "what happens to the subject, when they have a feeling of pastness associated with their memory, is that they experience one of the things represented by their memory" (Fernández, 2020, p. 294). A little later in the same paragraph, Fernández draws the comparison to the experience of blue. He says,

> looking at a blue wall, for example, I experience the fact that my perceptual state represents a certain reflectance property of the surface that I am looking at. I experience it in a qualitatively distinctive way; the way which is characteristic of seeing blue (Fernández, 2020, p. 294).

In the same way that we have genuine experiences of blue by representing reflectance properties, we have genuine experiences of pastness by representing causation.

We have a proposal. The feeling of pastness, a phenomenological feature of episodic memory, is grounded in the representation of causal origin *similar to how* our experience of blue is grounded in the representation of reflectance properties.

3. Representations of causal and temporal order are distinct

The goal of this section is to show that the self-referential account of PAST falls short. Appealing to the fact that memories represent their causal origins, even if causes as a matter of fact precede their effects, does not explain why memories possess a feeling of pastness.

Recall —the form of representationalism that Fernández adopts— the content of a memory provides information regarding that memory's phenomenal properties. In the color case, why would the visual representation of reflectance properties provide us with information about the phenomenal blueness that a subject enjoys in virtue of that representational state? We cannot simply appeal to the bare correlation between a state's having that representational content and its having a certain phenomenal property. That simply restates the explanandum by repeating the fact that states with this content have a certain phenomenology. That isn't an informative explanation. However, color objectivists (Byrne & Hilbert, 2003; Dretske, 1997; Tye, 2000) have an explanation. For something to be represented by our visual system as having a particular reflectance property just is for it be represented as being blue, because the property of blueness just is the property of having a particular type of reflectance property. The informativeness of the account depends on a metaphysical claim about the property picked out by 'is blue' and 'has such and such a reflectance property'.

While there is a debate over the metaphysics of color, the analogous debate doesn't exist with regards to causation and time.[4] Causation and time are closely related to one another, yet, they are not identical. Fernández might appeal to a weaker claim. Causal order is not identical to temporal order, but rather, causal order entails temporal order. Call this *the temporal priority principle* (Rankin & McCormack, 2013) according to which causes precede their effects (either contingently in our world or by necessity). Therefore, memory's representing causal origin entails something about temporal order, and this entailed content explains the temporal phenomenology.

Unfortunately, more would still have to be said about this. In general, we do not seem to have phenomenal experiences of all of the entailments of what our experiences represent. Furthermore, there are empirical reasons for doubting that the metaphysical explanation given by Fernández works. If Fernández is correct, then PAST is explained by the causal content of memory as a result of the metaphysical connection between causation and time. Therefore, if it is the metaphysics that is explaining this phenomenological fact, then if we have other experiences of causation, then we should also find that they will have the corresponding temporal phenomenology. The same metaphysical considerations would hold. However, that is not what we find in the perceptual cases that I will describe below. Our perceptual experiences can violate

---

4   Exceptions would be Reichenbach (1957) and Grunbaum (1967). See (Earman, 1972) for criticisms.

the temporal priority principle. If the representation of causation in memory explains PAST, then Fernandez will have to appeal to something other than the metaphysics of time and causation to make his point.

Let's turn our attention to perception to see how experiences of causal and temporal order come apart. It's widely accepted that we are capable of perceiving causation (Michotte, 1955; Rolfs, Dambacher & Cavanagh, 2013; Scholl & Tremoulet, 2000) as well as time (Buhusi & Meck, 2005). It's also becoming clear that our perception of time and causation influence each other.

Consider how the perceived timing of events serves as a cue for causation. In classic Michotte-style launching experiments, an object will be seen to move across a screen, it will come into contact with another object, and then the second object will begin to move along the same trajectory as the first object. If the objects move smoothly, their trajectories line up in the correct way, and if there is no delay between when the first object comes in contact with the second and the second object's movement, people will reliably perceive these sequences as *launching events* in which the first object causes the second to move (figure 1a). However, if there is a delay between when the objects come into contact and the movement of the second object, then the sense of causation goes away (figure 1b) (Michotte, 1955).
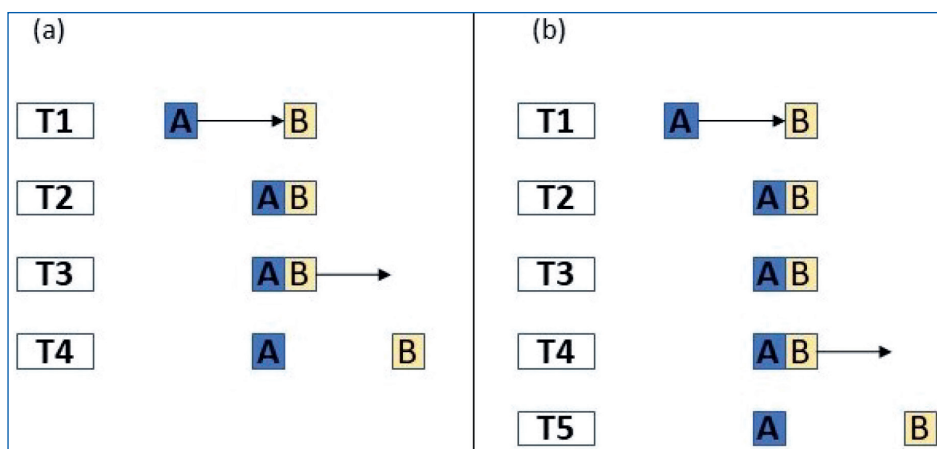
Figure 1
Michotte-style launching displays

(a) subjects perceive this interaction as a causal one when there is no delay; (b) the extended delay between A and B's movements eliminates the sense of causation.

The timing of crossmodal cues also influences our perception of causation. In the *bounce-stream illusion* (Sekuler, Sekuler, & Lau, 1997), subjects are presented with two lines emerging from the top-left and top-right of the screen heading on a diagonal, intersecting in the middle, and continuing on. If only the visual stimulus is presented,

subjects will perceive these two lines streaming past one another and continuing on straight paths (figure 2a). However, if an auditory click is presented when the two streams touch in the middle of the screen, then subjects perceive the two lines as colliding and bouncing off of each other. The result of this apparent causal interaction is that the lines do not proceed along straight paths, but bounce off of each other with the line from the top-left ending up in the bottom-left and vice versa (figure 2b). Perceived timing in launching experiments and the stream-bounce illusion serve as a cue for the perception of causation.
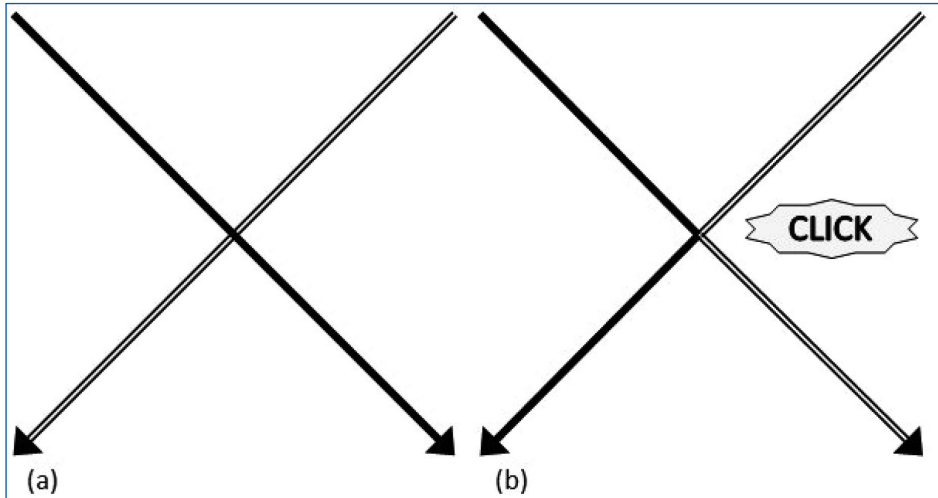
Figure 2
The bounce-stream illusion

The single and double lines represent perceived trajectories. Actual stimuli did not differ in this way. (a) in the absence of the sound, the lines appear to continue along straight paths. (b) with the properly timed sound the lines appeared to bounce.

More recent findings show that apparent causal relationships can influence our perception of time. In *temporal reordering cases* (Bechlivanidis & Lagnado, 2013; 2016), by manipulating causal cues, subjects can be induced to perceive a reversal of temporal order that matches the perceived causal order. In a modified Michotte-Style experiment, three objects, A, B, and C interacted. In the veridical causal condition, object A would move into contact with B, then B would move into contact with C, and then C would move. Subjects reliably perceived the correct causal and temporal order in this condition. The critical condition for their experiment was a modification of this standard sequence in which C would begin to move before B would —i.e. the effect would occur before the cause. In this sequence, subjects reliably misperceived the temporal order of events in a way that respects the temporal priority principle —i.e. they perceived the temporal sequence as A moved, then B, then C. In another condition, the relative timing

of B and C was kept the same as the previous condition in which C moved before B, but A was removed from the display. In this case, where the cue for the entire causal sequence is removed, subjects accurately perceived the temporal order of B and C's movements. Cues regarding causal interaction influenced the perception of temporal order.

Temporal binding effects also show how apparent causal relations can influence temporal perception. In the temporal binding effect, two events at different moments in time are perceived as occurring closer in time than they in fact did provided that the subject has some reason for perceiving one of these events as the cause of the other. A standard experimental design exhibiting this effect involves a subject pressing a button and then a tone is played. The subject will reliably perceive the button press and the tone as being closer in time than they in fact were. However, if the subject is given cues indicating that these events are not causally related, then the temporal binding effect is eliminated (Chen & Vroomen, 2013; Suzuki, Lush, Seth & Roseboom, 2019).[5] Causal perception once again seems to influence temporal perception.

Everything said so far is compatible with a representation of causation in memory being "informative" in some sense regarding the experience of time. However, the amenable story ends here. The temporal binding effect can be exploited to produce experiences that violate the temporal priority principle.

In a now classic study by Stetson et al. (2006), subjects were asked to press a button and then after a variable delay, with an average length of 35ms, a flash of light would appear on the screen in front of them. Subjects reliably perceived the flash of light as occurring after the button press. In a second block of trials, the experimenters inserted an extended delay of approximately 135ms between the button press and the flash of light. An interesting finding with experimental situations like this is that there is a progressive adaptation effect that leads to a stronger temporal binding effect after repeated trials. The relevant finding came in the third block of trials. After adaptation the extended delay was removed and the flash of light once again appeared approximately 35ms after the button presses. Even though the timing of the stimuli in this third block of trials was identical to the timing from the first block, subjects reliably perceived the flash of light as occurring prior to their button presses. However, recall that the temporal binding effect only took hold given that the subject had a sense of the causal interaction between their button presses and the flash of light. In this study, which has been replicated in (Cunningham, Billock, & Tsou, 2001; Heron et al., 2009), subjects had a very peculiar perceptual experience. The apparent causal and temporal order of their actions and the effects of those actions violated the temporal priority principle. Effects seemed to occur prior to their causes!

---

5   Hoerl, Lorimer, McCormack, Lagnado, Blakey, Tecwyn & Buehner, (2020) have argued that the temporal binding effect results from a top-down influence of causal belief on perception. However, it is unclear whether this explanation can account for those cases described in the next paragraph that exploit adaptation effects.

In perception, the experience of causal order and the experience of temporal order come apart. A perceptual representation of causal order does not, by itself, ground an experience of pastness or of temporal order. If Fernández wants to account for PAST by appealing to the causal content in memory, then appealing to the metaphysics of time won't work. The same considerations would apply to perception, yet in perception we see that the temporal priority principle can be violated. If the explanation of PAST developed by Fernández is to succeed, then he must point to something about memory, and not about the metaphysics of time and causation, that would show why his explanation succeeds. Without this additional explanation, we do not have an account of PAST.

However, another option is possible. We take the representation of causation and time to be defeasible cues for one another. The representation of causation in memory could be used as a cue for the representation of temporal location or pastness in memory. An explicit representation of time would provide us with an explanation for our feeling of pastness. Yet, if Fernández is correct, then no version of temporal location can successfully account for PAST. In the next section, I'll sketch out a version of temporal location that avoids the worries raised by Fernández and allows us to account for PAST.

4. Path dependent representations

The goal of this section is to spell out an alternative subject-dependent account of temporal localization −*path-dependent representations.* To motivate the account, it will be helpful to begin with spatial representation.

The standard division in the spatial literature is between allocentric and egocentric representations. Allocentric representations pick out spatial locations in a way that makes no necessary reference to an agent. A map can indicate the location of my office in Sheffield without any reference to where I am in relation to my office. This is the spatial analog of Fernández's subject-independent account of temporal localization.

Egocentric representations, on the other hand, pick out locations relative to an agent. However, types of egocentric representation can differ depending on how they specify locations relative to an agent. Here, I will distinguish between two types of egocentric representation. The first are *target representations.* Consider my office again. I might represent the office relative to my location by describing a single vector that picks out that location "as the crow flies". For instance, I might represent my office *as being 1.3 miles North of here.* This form of egocentric representation is the spatial analog of Fernández's subject-dependent account of temporal localization.

The other form of egocentric representation are what I call *path-dependent representations.* Path-dependent representations specify locations by the paths that

would be taken to arrive at a target.[6] This sort of localization should be familiar to someone that has moved to a new city. You might know how to get to your office by following a specific path. I might know that my office is 30 minutes away if you follow this meandering street then turn left at the big intersection, then turn right below the bridge, etc. You can know where your office is in this way without knowing anything about the absolute location of your office or knowing where you stand relative to the office "as the crow flies".

My suggestion is that we should think of temporal representations of pastness as path-dependent representations. If we do this, then we will have an account of temporal representation that avoids the problems Fernández raised against the temporal location accounts of PAST.

One initial objection to applying path representations to time is that we have multiple dimensions of space that we can travel through. Therefore, different paths in space to a single target can differ in length. However, time is a single dimension. There are no choices of paths between temporal points. As a result, path-dependent representations of time devolve into Fernández's subject-dependent representations, since our only path through time is "as the crow flies". However, it's not so clear that this is true. To see why it is helpful to think about clocks and then we'll turn to time travel.[7]

Suppose someone travels from London to New York City. They carry a clock while they do this. What does their clock measure? What does the clock of their friend who remained in London measure? In a classical spacetime, the clocks simply measure the absolute amount of time that passes between the person leaving London and arriving in NYC. All observers, and clocks, should agree on how much time passed over this interval. In a relativistic spacetime, like our world, the situation is different. There is no absolute time. Instead, our clocks measure the temporal distance along the distinct spacetime worldlines traced by the traveler and their sedentary friend. As a result, the two clocks will disagree on how much time has passed since relativistic effects change the temporal length of these worldlines. This isn't just a theoretical result. It's something that has been observed and measured.[8] A regular stopwatch already measures the length of a temporal path and not absolute time.

Consider our time travel case again. I remember struggling to swim at some point in my past. That event took place in 1991. It is now 2021. I step into a time machine and emerge in 1989. I still remember this event accurately and still have a feeling of pastness associated with my memory. How do we account for this feeling of pastness? It's useful to distinguish between two notions of time in these scenarios. There is *external time* and *personal time* (Lewis, 1976). External time locates events in time via an egocentric

---

6   These representations are similar to the action guiding egocentric representations that we find in of P. F. Evans (1982), Grush (2000) & Strawson (1964).

7   The view developed here is indebted to Lewis (1976).

8   The details for this go beyond the space I have here. For accessible explanations see (Callender, 2017).

target framework. In 1989, the remembered event is two years in the future. However, *personal time* is the time measured by the time traveler's watch – it's a measure of time along *my personal* worldline. As I enter my time machine and travel backwards in (external) time, from my perspective time and causation continue to move forward.

When I remember my struggling to swim in my past, I do not represent the pastness of this event in external time. Rather, I represent the pastness of the event in a path-dependent way —it is *my personal causally continuous past.* When I emerge in 1989, the remembered event will be in the past of my worldline. It will be represented as being in that temporal past that was causally relevant to my current memory (perhaps through a memory trace). There is no inaccuracy here. We have a way of temporally locating events that seems to capture PAST. Given the representationalism that Fernández endorses, the fact that our memories represent events as being in our path-dependent past would explain why memories have the phenomenology of pastness.

The suggestion here isn't that we use (internal) clocks to represent the temporal location of remembered events. Rather, we use a variety of cues to locate remembered events in a path-dependent representation of time. The causal content of memory could be a cue for this sort of representation. That a perceptual episode caused our current memory is reason for representing that that perceptual episode was in the (path-dependent) past since the causal influence from that perceptual episode moves forward along our worldline.

Fernández could maintain the vast majority of his self-referential account and still provide an explanation of PAST if he includes path-dependent temporal content as a component of his theory. Otherwise, he would have to either abandon attempts at explaining PAST, since causal content alone does not suffice, or he has to provide a reason for thinking that the representation of causal order in memory has a different connection to time and temporal phenomenology than it does in perception.

However, this account of temporal representation is also compatible with constructivist approaches to memory (De Brigard, 2014; Michaelian, 2016). If, in the act of recollection, we utilize a variety of mental representations to construct a representation of a past event, then path-dependent temporal representations may be an element in this process. Information about causal origin, along with other information, may serve as a cue for forming the appropriate path-dependent temporal representation in the act of recollection.

Finally, I'm not suggesting that in remembering we have a theoretical grasp of path-dependent temporal representation. Rather, we utilize path-dependent representations of time to locate events. If we accept anti-individualism about content (Burge, 1979), i.e. that what we represent is determined by the world, and we accept that temporal properties are path-dependent, then our temporal representations will latch onto path-dependent temporal properties. In the same way that we can use clocks that measure path-dependent time without understanding relativistic physics, the same is true for the mental representation of time.

5. Conclusion

The target of this paper was whether the self-referential theory could account for PAST. I gave reasons for why it can't. It can, however, be augmented with path-dependent temporal representations while retaining the connection between causation and temporal order that Fernández was appealing to. The difference is that it involves explicit temporal representation.

An important question has been lingering beneath the surface of this paper. Memories are complex and what information is accessed at the moment of recall often depends on the context and purpose of recall (Michaelian, 2016). Much of this information is informative regarding the timing of events. Things like familiarity, visual cues, location, etc. can be used to infer something about time. Phenomenology is also difficult to report. It may be the case that the distinctive phenomenology of memory, what we described as PAST, might not be a single phenomenal experience. Rather, remembering may have various phenomenal features, some temporal, some causal, some familiarity-based (Perrin, Michaelian & Sant'Anna, 2020), and we simply describe these phenomenal features in temporal terms due to our familiarity with temporal discourse. This paper hasn't engaged with this question. However, if we want to account for PAST, something more than representations of causation is needed. I've articulated an account of path-dependent temporal representations that can help.

Acknowledgements

References

Bechlivanidis, C. & Lagnado, D. A. (2013). Does the "why" tell us the "when"? *Psychological Science, 24*(8), 1563–1572. https://doi.org/10.1177/0956797613476046

Bechlivanidis, C. & Lagnado, D. A. (2016). Time reordered: causal perception guides the interpretation of temporal order. *Cognition, 146*, 58–66. https://doi.org/10.1016/j.cognition.2015.09.001

Boyle, A. (2020). Remembering events and representing time. *Synthese*, 1–20. https://doi.org/10.1007/s11229-020-02896-6

Buhusi, C. V. & Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nature Reviews. Neuroscience, 6*(10), 755–765. https://doi.org/10.1038/nrn1764

Burge, T. (1979). Individualism and the mental. *Midwest Studies in Philosophy, 4*(1), 73–122. https://doi.org/10.1111/j.1475-4975.1979.tb00374.x

Byrne, A. & Hilbert, D. R. (2003). Color realism and color science. *Behavioral and Brain Sciences, 26*(1), 3–21. https://doi.org/10.1017/S0140525X03000013

Callender, C. (2017). *What makes time special?* Oxford University Press. https://doi.org/10.1093/oso/9780198797302.001.0001

Chen, L., & Vroomen, J. (2013). Intersensory binding across space and time: a tutorial review. *Attention, Perception, & Psychophysics, 75*(5), 790–811. https://doi.org/10.3758/s13414-013-0475-4

Cunningham, D. W., Billock, V. A. & Tsou, B. H. (2001). Sensorimotor adaptation to violations of temporal contiguity. *Psychological Science, 12*(6), 532–535. https://doi.org/10.1111/1467-9280.d01-17

De Brigard, F. (2014). Is memory for remembering? Recollection as a form of episodic hypothetical thinking. *Synthese, 191*(2), 1–31. https://doi.org/10.1007/s11229-013-0247-7

Dretske, F. (1997). *Naturalizing the mind*. MIT Press. https://doi.org/10.7551/mitpress/4872.001.0001

Earman, J. (1972). Notes on the causal theory of time. *Synthese, 24*(1), 74–86. https://doi.org/10.1007/BF00540142

Evans, G. (1982). *The varieties of reference*. Oxford University Press.

Fernández, J. (2019). *Memory: a self-referential account*. Oxford University Press. https://doi.org/10.1093/oso/9780190073008.001.0001

Fernández, J. (2020). Self-referential memory and mental time travel. *Review of Philosophy and Psychology, 11*(2), 283–300. https://doi.org/10.1007/s13164-019-00453-w

Grunbaum, A. (1967). *Modern science and Zeno's paradoxes*. George Allen & Unwin.

Grush, R. (2000). Self, world and space: the meaning and mechanisms of ego-and allocentric spatial representation. *Brain and Mind, 1*(1), 59–92. https://doi.org/10.1023/A:1010039705798

Heron, J., Hanson, J. V. M. & Whitaker, D. (2009). Effect before cause: supramodal recalibration of sensorimotor timing. *PLOS ONE, 4*(11), e7681. https://doi.org/10.1371/journal.pone.0007681

Hoerl, C., Lorimer, S., McCormack, T., Lagnado, D. A., Blakey, E., Tecwyn, E. C. & Buehner, M. J. (2020). Temporal binding, causation, and agency: developing a new theoretical framework. *Cognitive Science, 44*(5), e12843. https://doi.org/10.1111/cogs.12843

Lewis, D. (1976). The paradoxes of time travel. *American Philosophical Quarterly, 13*(2), 145–152.

Michaelian, K. (2016). *Mental time travel: episodic memory and our knowledge of the personal past*. MIT Press. https://doi.org/10.7551/mitpress/10591.001.0001

Michotte, A. (1955). La Perception de la causalité. *Les Etudes Philosophiques, 10*(2), 308–309.

Perrin, D. (2018). A case for procedural causality in episodic recollection. In K. Michaelian, D. Debus & D. Perrin (Eds.), *New Direction in the Philosophy of Memory* (pp. 33–51). Routledge. https://doi.org/10.4324/9781315159591-3

Perrin, D., Michaelian, K. & Sant'Anna, A. (2020). The phenomenology of remembering is an epistemic feeling. *Frontiers in Psychology, 11,* 1531. https://doi.org/10.3389/fpsyg.2020.01531

Rankin, M. L. & McCormack, T. (2013). The temporal priority principle: at what age does this develop? *Frontiers in Psychology, 4.* https://doi.org/10.3389/fpsyg.2013.00178

Reichenbach, H. (1957). *The philosophy of space and time* (1st English Ed. edition). Dover Publications.

Rolfs, M., Dambacher, M. & Cavanagh, P. (2013). Visual adaptation of the perception of causality. *Current Biology, 23*(3), 250–254. https://doi.org/10.1016/j.cub.2012.12.017

Scholl, B. J. & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences, 4*(8), 299–309. https://doi.org/10.1016/S1364-6613(00)01506-0

Sekuler, R., Sekuler, A. B. & Lau, R. (1997). Sound alters visual motion perception. *Nature, 385*(6614), 308–308. https://doi.org/10.1038/385308a0

Stetson, C., Cui, X., Montague, P. R. & Eagleman, D. M. (2006). Motor-sensory recalibration leads to an illusory reversal of action and sensation. *Neuron, 51*(5), 651–659. https://doi.org/10.1016/j.neuron.2006.08.006

Strawson, P. F. (1964). *Individuals: an essay in descriptive metaphysics.* Routledge.

Suzuki, K., Lush, P., Seth, A. K. & Roseboom, W. (2019). Intentional binding without intentional action. *Psychological Science, 30*(6), 842–853. https://doi.org/10.1177/0956797619842191

Tye, M. (2000). *Consciousness, color, and content.* MIT Press. https://doi.org/10.7551/mitpress/2110.001.0001

Immunity to error through misidentification and the functionalist, self-reflexive account of episodic memory*

Steven James

West Chester University of Pennsylvania, Pennsylvania, United States
E-mail: sjames@wcupa.edu

Abstract: Fernández (2019) offers an account of the nature of episodic memory that marries two core ideas: (i) role-functionalism about episodic memory, and (ii) self-reflexive mnemonic content. One payoff of this view is that episodic memory judgments are immune to error through misidentification. Fernández takes this to reveal something important about the nature of one's self-awareness in memory and our first-person conception of ourselves. However, once one sees *why* such judgments are immune in this way, according to the proposed account, the fact that they are immune becomes moot. While technically immune to error through misidentification, episodic memory judgments are not grounded in a way such that they have any interesting epistemological import for the subject (in contrast to other paradigms of such judgments), and any insights about our self-awareness and self-conception are directly derivable from the metaphysics of memory content alone.

Keywords: immunity to error through misidentification, memory judgments, episodic memory, self-awareness, metaphysics of memory

---

Inmunidad al error de identificación errónea y la teoría funcionalista y autorreflexiva de la memoria episódica

Resumen: Fernández (2019) ofrece una teoría de la naturaleza de la memoria episódica que reúne dos ideas centrales: (i) el funcionalismo sobre la memoria episódica y (ii) el contenido mnemónico autorreflexivo. Una ventaja de esta teoría es que los juicios de memoria episódica son inmunes al error por identificación errónea. Según Fernández, esto revela algo importante sobre la naturaleza de la conciencia de sí mismo en la memoria y nuestra concepción en primera persona de nosotros mismos. Sin embargo, una vez que uno ve por qué tales juicios son inmunes, según el relato propuesto, el hecho de que sean inmunes se vuelve discutible. Si bien técnicamente son inmunes al error a través de la identificación errónea, los juicios de memoria episódica no se basan en una forma tal que tengan una importancia epistemológica interesante para el sujeto (en contraste con otros paradigmas de tales juicios), ya que cualquier conocimiento sobre nuestra autoconciencia y autoconcepción se derivan directa y únicamente de la metafísica del contenido de la memoria.

Palabras clave: inmunidad al error de identificación errónea, juicios de memoria, memoria episódica, auto-conciencia, metafísica de la memoria

Steven James

Is associate professor at the West Chester University of Pennsylvania. His specializations are analytic epistemology, philosophy of mind, cognitive science, and philosophy of language. He is currently completing a book-length project on the nature of remembering things.

ORCID: 0000-0002-8370-8746

Fernández (2019) offers a novel, interesting, and quite clever account of the nature of episodic memory that marries two core ideas: (i) role-functionalism about episodic memory —a mental state counts as a memory in virtue of playing a particular functional role in the cognitive system of the subject (chapter 2), and (ii) self-reflexive mnemonic content— the contents of episodic memories are quite 'thick' in that they are necessarily constituted by representations of *all* of the following: (a) the memory itself, (b) a past (objective) fact, (c) one of the remembering subject's past perceptual experiences, and (d) a causal relationship between that past experience and the present memory (chapter 3). One payoff of this view is that judgments on the basis of genuine episodic memories are, despite appearances, necessarily free of a certain kind of error —those involving the misidentification of the self in memory. In other words, they are immune to error through misidentification, hereafter 'IEM' (chapter 6). Fernández takes this surprising result to reveal something important about one's self-awareness in memory (p. 143) and our first-person conception of ourselves (p. 169). However, once one sees *why* such judgments are IEM, according to the proposed account, the fact that they are IEM becomes considerably less interesting. Episodic memory, while technically IEM, does not *meaningfully* ground judgments that have any special epistemological import for the subject (in contrast to other paradigms of IEM judgments). And any insights about self-awareness and self-conception are directly derivable from the metaphysics of memory content alone.

The article is structured as follows. In section 1, I introduce the notion of IEM, and review some reasons that states with such status could be theoretically important. In section 2, I present a case that common sense, and Fernández's own functionalist account of episodic memory, would suggest is an instance of episodic memory that fails to be IEM in the relevant way. I then explain why it would not be surprising for episodic memory to fail to be IEM. In section 3, I explain why Fernández is, nevertheless, in a position to reject this conclusion by appealing to his account of episodic memory content in conjunction with the particular formulation of the IEM thesis he has adopted. In section 4, I show that this way of establishing that episodic memory is IEM strips the claim of its theoretical interest.

1. Immunity to error through misidentification (IEM)

Despite the well-known fallibility of human cognition, many philosophers are tempted by the thought that the scope of our ability to be mistaken is somehow limited; there are certain domains in which one's judgments have some special kind of epistemic status. Of particular interest here is the claim that there are certain kinds of judgments which avoid the possibility of 'mistaken identity'. In other words, there are judgments in which it is impossible to form, 'a false belief because of a misidentification of the person or object about whom one made the judgment' (Prosser and Recanati, 2012, p. x).

In the literature, such judgments are labeled immune to error through misidentification (IEM) and contrasted with judgments which *are* subject to error through misidentification. The least controversial candidates are those involving first-person judgments about oneself on the basis of occurrent conscious experiences, e.g. introspective judgments about one's current conscious mental states, proprioceptive judgments. Borrowing again from Prosser and Recanati (2012, p. ix), two exemplars are, "'I have a headache' (where the judgment is made on the grounds that one feels one's head aching) and 'my legs are crossed' (where the judgment is made on the basis of proprioception). In the first case, it is said, I could not be mistaken about whose headache I was aware of; and in the second kind of case I could not be mistaken about whose legs I felt to be crossed." In contrast, judgments like, 'Bill was rude at the party' (on the basis of observing his behavior) are subject to error through misidentification because it is possible for one to mistake someone else for Bill.

Following Prosser (2012, p. 161), it seems that what makes a judgment susceptible to error is a certain kind of structure: (i) some object (*a*) is the causal source of information that leads to (motivates/justifies, etc.) the application of a predicate (*F*). (ii) The predicate (*F*) gets applied to an object (*b*). And error arises when $a \neq b$. So, one should expect a formula for identifying judgments that are immune to such errors. Find those judgments, should there be any, in which the causal source that leads to the application of the predicate, and the object to which the predicate gets applied cannot fail to be identical.[1]

Finding such types of judgments could be interesting for a variety of reasons. It might inform our understanding of the structure of justification —IEM judgments may have a kind of epistemological priority (Evans, 1982, pp. 181-182). It might inform our understanding of the semantics of indexicals and demonstratives (Evans, 1982; Recanati, 2007). It might inform our understanding of the structure of self-knowledge (Evans, 1982; Ismael, 2012; Merlo, 2017). And most importantly for present purposes, it might reveal something important about our awareness of ourselves in memory, and our conception of the self (Evans, 1982; Fernández, 2019).[2] As Fernández sees it, '[i]f memory judgments are IEM, then (...) in memory, one is aware of the subject who is remembered to have instantiated such-and-such properties as being oneself' (2019, p. 143). More precisely, the IEM status of memory judgments would reveal,

> the fact that, in memory, we are presented to ourselves (...) as the bearers of extrinsic properties which were perceived in the past; properties such as occupying a certain spatial position or having a particular size relative to that of another object (...) [with a final lesson being] that our first-person conception of

---

1   Cappelen & Dever (2013, pp. 130-33) make a reasonably compelling case that there is *no* philosophically interesting phenomenon of immunity to error through misidentification (building on cases described in Higginbotham [2010]), but I'll grant that there are at least some interesting cases of IEM judgments here.

2   While representative of reasons for interest in IEM, this list is by no means exhaustive.

ourselves does not only include the fact that we are thinking things, or bearers of mental properties, as Descartes may have suggested (...) And our first-person conception of ourselves does not only include the fact that we are bearers of physical properties such as being extended in space, as Evans suggests (...) Our first person conception of ourselves also includes the fact that we are the bearers of temporal properties. Our First-person conception of ourselves in other words, is the conception of an object which has a history (Fernández, 2019, pp. 169-70).

If Fernández is right, determining whether memory judgments are immune to error through misidentification should have profound implications, and it is quite understandable why the thesis has enjoyed considerable popularity.[3] Nevertheless, the thought that episodic memory might have some special epistemic status, that misidentifications in it might be *impossible,* is at least somewhat surprising for the reasons I turn to next.

2. Why episodic memory doesn't seem to be IEM

Let us adopt Fernández's *prima facie* plausible functionalist account of episodic memory according to which (episodically) remembering a fact consists in having a mental image that plays the mnemonic role for that fact in the subject.

> [F]or any subject *S* and proposition *p, S* remembers that *p* just in case *S* has some mental image *i* such that *i* tends to cause in *S* a disposition to believe both that *p* and that *S* experienced that *p,* and *i* tends to be caused in *S* by having experienced that *p* (Fernández, 2019, p. 49).

Now consider the following fictional case adapted from the cartoon *Bob's Burgers.*[4]

> Aunt Gayle: It was my second sophomore year in junior college, and I took a trip to New York City with...my sister Linda...We were in the denim district...when an elegant woman walked by. I looked up, she looked at me, and she gave me a wink and the finger guns... And that woman was Delta Burke...It was the briefest of gestures, but it meant so much to me. A Designing Woman. That moment helped design this woman.

---

3   It is endorsed, in one form or another, by Bermúdez (2012; 2013), Evans (1982), Hamilton (2007), McCarroll (2018), Recanati (2007), and Shoemaker (1970) among others, though naturally, not necessarily for the reasons Fernández provides.

4   Adapted from *Bob's Burgers,* Season 7, Episode 22: 'Into the Mild' (2017). While the case is fictional, it depicts a familiar phenomenon. See discussions of it and related phenomena in the empirical literature in, e.g. Brown, Croft Caderao, Fields & Marsh (2015); Pasupathi & Wainryb (2018); Pillemer, Steiner, Kuwabara, Thomsen & Svob, (2015); Reese & Brown (2000); Ross & Ward, (1996) and Sheen, Kemp & Rubin, (2001). Michaelian (2020) briefly discusses such cases but does not pursue them as part of his argument against the claim that episodic memory is IEM on the basis of the existence of observer memory.

Linda:	Unbelievable.
Gene:	Seriously, it's hard to believe Aunt Gayle was close enough to have "smelt-a" the Delta.
Linda:	Yeah, hard to believe because it didn't happen to her. It happened to me.
Louise:	Wait, Mom, are you saying that Aunt Gayle stole your Delta Burke wink and finger gun story?
Linda:	Yeah, it was me who got winked and gunned. We had a connection. I'll never forget her eyes.
Tina:	If it bothers you, you should say something, Mom.
Linda:	No, no, it's fine. It means so much to her, let her have it. I've got a good life.

Suppose that Aunt Gayle is sincere in her recounting of the story. She is telling it on the basis of her memory of the trip that she and her sister took, and her belief that Ms. Burke gave her a 'wink and the finger guns' is based on a mental image as of it happening just as she described. Such a case satisfies FTM. The subject has a kind of mental image that tends to cause in her a disposition to believe both that *p* and that she experienced that *p*. And such an image is the kind of thing that tends to be caused by such an event happening.[5] In short, Gayle's mental image plays the mnemonic role for her, and so is an episodic memory that *p*.

Now, suppose also that Linda is right. While Gayle was present and saw the 'wink and the finger guns' happen just as retold, she was wrong about to *whom* it happened; it happened to Linda. In such a case, Gayle's belief appears to be an *exemplar* of a judgment that is subject to error through misidentification. It gets everything right, and is based on a memory that gets everything right, save the identification of who the event happened to (i.e. who exemplified the property in question).[6]

Such errors are familiar and perhaps happen more frequently than we recognize. On reflection, this should not be surprising, given what we know about the mechanisms by which memories are formed, modified over time, and reconstructed (often repeatedly).[7] Memory processes simply aren't suited to provide *immunity* to errors in general; identification of subjects in memory would be quite an outlier if it *were immune* in the way characteristic of IEM.[8]

---

5   This assumes that Gayle is not somehow systematically disposed to delusions. Compare Fernández's discussion of Korsakoff's patients (2019, chapter 2).

6   Compare, for example, Wittgenstein's (1958, pp. 66-7) example of mistaking a neighbor's broken arm for one's own on the basis of accurately seeing said arm after a car accident. In that case, the subject would be right in believing that someone's arm exemplified the property of being broken, but wrong in believing that they were that someone.

7   See, e.g. Schacter & Addis (2007) for a representative model of memory processing.

8   Michaelian (2020) makes effectively the same point (section 1)

Moreover, when you compare beliefs grounded in episodic memory to other judgments that are plausibly IEM, something is crucially different. Here is, for example, Jim Pryor on pain experiences:

> Beliefs of the form *I have a pain* are surely [IEM] when they are believed on the basis of introspective grounds. How could I possibly know on those kinds of grounds that *someone* has a pain, but be wrong in believing that it's me who feels a pain? *One can't be introspectively aware of a pain without thereby feeling that pain, oneself* (Pryor, 1999, p. 283; emphasis added).[9]

And here is Giovanni Merlo on the feeling of thirst:

> [I]n judging *that I'm thirsty*, I am not completely sure to be right, but I can 'rest assured' that
>
> I am not committing an error through misidentification (it would hardly make sense *for me* to wonder whether I'm mistaking someone else's thirst for mine) (Merlo, 2017, pp. 613-4), drawing on Wittgenstein (1958, pp. 66-7).

In such cases, and in contrast to 'ordinary judgments', there is a sense in which the author of the judgment *can rule out* the possibility that their judgment involves a case of misidentification.[10] Jim can rule out the error because introspectively being aware of a pain involves feeling that pain oneself. Giovanni can rule out the possibility of error because, in making the judgment about the thirst, there is a sense in which it is *nonsensical* for him to wonder whether he's mistaking whose sensation it is.[11]

Judgments on the basis of episodic memory are not like this. Gayle's awareness of the wink as happening to her doesn't make it have happened to her and, far from being able to rule out an error of misidentification, wondering whether one's memory involves such an error is eminently sensible.[12] Nevertheless, Fernández argues that despite such appearances, episodic memory really is IEM, and I turn to this next.

3. IEM and mnemonic content

Fernández's argument hinges on two choices. First, he follows Shoemaker in characterizing IEM in the context of episodic memory as follows:

---

9   Shoemaker (1968, pp. 563-4) goes so far as to claim that 'In being aware that one feels pain one is, tautologically, aware not simply that the attribute *feels pain* is instantiated, but that it is instantiated in *oneself.*'

10   Merlo (2017, pp. 613-4). Cappelen & Dever suggest that such ruling out could be done *a priori* (2013, pp. 131)

11   Here again, I'm setting aside the argument in Cappelen & Dever (2013, pp. 130-133).

12   Such wondering occupies much of the remainder of the plot of the episode of the cartoon, and is integral to much of the research cited in note 4.

> My memory report could of course be mistaken, for one can misremember such incidents, but it could not be the case that I have a *full and accurate memory of the past incident* but am mistaken in thinking that the person I remember [exemplifying the property] was myself (Shoemaker, 1970, pp. 269-70, emphasis added).

The 'full and accurate memory' specification is crucial to Fernández's understanding of IEM and what it takes for a judgment to fail to be IEM. Here is his definition of IEM (2019, p. 142):

> For any property P and grounds G:
>
> If S judges that S has P on the basis of G, then that judgment is IEM relative to G iff it is impossible that there is a subject S* such that:
>
> G represents S* as having P.
>
> G is fully accurate.
>
> S mistakenly thinks that S is identical with S*.
>
> S's judgment that they have P is false because of (iii).[13]

Defined in this way, whether a judgment is IEM or not, is not merely a matter of whether or not it is impossible to form a false belief because of a misidentification of the person or object about whom one made the judgment;[14] and a failure of IEM does not necessarily happen when, in forming a judgment that is otherwise correct, someone mistakes something for something else.[15] For Fernández, a failure of IEM requires a misidentification on the basis of grounds that are *accurate full stop.* The result is that whether episodic memory judgments are IEM crucially depends on the nature of the relevant grounds, i.e. on the nature of the content of episodic memory.

The second key choice is Fernández's decision to adopt a reflexive account of episodic memory content.[16] It is as follows:

> Reflexive View
>
> For any subject S, memory M and proposition q:
>
> If S has M and S would express M by saying that they remember that q, then there is a perceptual experience P that S would express by saying that they perceive q, such that the content of M is the proposition {W: In W, M is caused by S having perceived that q through P}[17]

---

13   For presentational purposes, I have modified Fernández's formulation in ways that do not change the account.

14   Pace the earlier mentioned Prosser and Recanati (2012, p. x)

15   Pace Merlo (2017, p. 605)

16   This choice is, of course, eminently understandable, given that he spent chapter 3 motivating said account.

17   Fernández (2019, p. 79)

Defining mnemonic content in this way entails that one's memory (i.e. 'G' in the above formulation of IEM) is only fully accurate if the remembering subject is identical to the person or object about whom the memory judgment is made. Prosser's formula is thus satisfied as it is impossible for the causal source that leads to the application of the predicate and the object to which the predicate gets applied to fail to be identical. In other words, there is no room for misidentification. So, while subjects may have memory-related judgments that involve misidentification as described in section 2, such judgments do not, technically, fail to be IEM because they are not formed on the basis of grounds that are accurate full stop.

4. Re-examining the significance of IEM

Once one grants Fernández's definitions of IEM and mnemonic content, it follows, effectively by definition, that episodic memory judgments are immune to error through misidentification.[18] In this section, I conclude by briefly arguing that this result is not as interesting as Fernández suggests.

First, while it may be that memory judgments are technically IEM, they are still importantly different from other paradigmatic cases of IEM judgments. For example, as we saw above, in judging that one is in pain on the basis of introspection, one's awareness (and corresponding judgment) of the sensation at that moment *involves oneself having it at that moment*. It is this intimate connection that purportedly makes it 'nonsensical' to wonder whether it is oneself undergoing the experience, and that grants the author of the judgment some kind of epistemic assurance, i.e. that explains its being IEM.[19] In contrast, on the current proposal, it is the fact that any memory-related judgment that fails to be IEM will also fail to be based on 'fully accurate' grounds that explains its being IEM.[20] Fernández rightly, I think, argues that assessing the IEM status of episodic memory by examining quasi-memories would be unproductive because quasi-memories are *defined* in such a way that errors through misidentification *are possible* (2019, Chapter 6, section 3). However, whereas quasi-memories are defined such that errors through misidentification must be possible,

---

18 Considerations of space preclude an examination of Fernández's arguments in favor of a self-reflexive view of mnemonic content, but such views of mental content are by no means uncontroversial. See Michaelian (2020, section 7) and Tye (2009, p. 80) for distinct criticisms of the self-reflexive view of content for memory specifically. See e.g. Millar (1991), Soteriou (2000), and Tye (2009) for broader criticism of self-reflexive views of content. On the other hand, see, e.g. Fernández's earlier work (2006) in its favor, Horgan & Kriegel (2007) and Levine (2018).

19 See the above discussion of Merlo (2017), Pryor (1999, p. 283) and Shoemaker (1968) −note 9 and associated text. See also Cappelen & Dever (2013, pp. 130-133) for an argument that such questions are *not* nonsensical at all.

20 It is also worth noting that in other paradigmatic cases of IEM judgments, whether they are based on *fully accurate* grounds appears to be beside the point. For example, it may be that one is wrong that it is *pain* or *thirst* that one is feeling (i.e., one's representation may be inaccurate in various ways) and yet it may remain the case that such judgments are IEM.

Estud.filos n.º 64. Julio-diciembre de 2021 | pp. 189-200 | Universidad de Antioquia | ISSN 0121-3628 | ISSN-e 2256-358X

197

https://doi.org/10.17533/udea.ef.n64a10

Fernández endorses *definitions* of mnemonic content and IEM that make (the relevant) misidentifications *impossible.* While this enables him to draw the desired conclusion, it also rings hollow because it secures his desired conclusion by fiat and thereby trivializes the thesis.[21]

As a result, and more importantly, it is not clear that it is the fact that episodic memory judgments are IEM that reveals anything about our awareness of ourselves in memory, or our self-conception. On the present proposal, memory judgments are *only* IEM once one adopts a particular account of mnemonic content (RV). Importantly, according to said account, memories represent, among other things, past experiences had by the remembering subject and causal relations between those past experiences and the remembering subject's current representations. Thus, it is the account of memory content that delivers insight into how we are presented to ourselves in memory, and the nature of our self-conception as of being the bearers of temporal properties —objects with histories. It is not the fact that memory judgments are IEM. In sum, the means by which Fernández's proposal secures the claim that episodic memory judgments are IEM strips that quite surprising result, and potentially interesting claim, of its theoretical import.

References

Behan, D. (1979). Locke on persons and personal identity. *Canadian Journal of Philosophy, 9,* 53–75. https://doi.org/10.1080/00455091.1979.10716236

Bermúdez, J. L. (2012). Memory judgments and immunity to error through misidentification. *Grazer Philosphische Studien, 84*(1), 123-142. https://doi.org/10.1163/9789401207904_007

Bermúdez, J. L. (2013). Immunity to error through misidentification and past-tense memory judgments. *Analysis, 73*(2), 211-220. https://doi.org/10.1093/analys/ant002

Brown, A. S., Croft Caderao, K., Fields, L. M. & Marsh, E. J. (2015). Borrowing personal memories. *Applied Cognitive Psychology, 29*(3), 471-477. https://doi.org/10.1002/acp.3130

Cappelen, H. & Dever, J. (2013). *The inessential indexical: on the philosophical insignificance of perspective and the first person.* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199686742.001.0001

---

21   A referee helpfully noted that this parallels the circularity of a so-called Lockean criterion for personal identity/persistence (whether Locke endorsed this view is questionable (see Behan [1979] cited in Olson [2019]). In short, one can secure the result that person A at *t2,* is identical with person B at *t1* only if A at *t2,* can remember an experience had by B at *t1,* by *assuming* that one can only remember one's own experiences. However, doing so makes the result is trivial. See, e.g. the discussions in Grice (1941, p. 344), Shoemaker (1959, p. 870); (1970, p. 270), and Parfit (1984, pp. 219-23). While Fernández does much to argue for his self-reflexive account of memory content, in the context of a distinct argument for the claim that episodic memory judgments are IEM, adopting the self-reflexive account amounts to begging the question in much the same way that assuming one can only remember one's own experiences begs the question of personal persistence.

Evans, G. (1982). *The varieties of reference.* Oxford University Press.

Fernández, J. (2006). The intentionality of memory. *Australasian Journal of Philosophy, 84,* 39-57. https://doi.org/10.1080/00048400600571695

Fernández, J. (2014). Memory and immunity to error through misidentification. *Review of Philosophy and Psychology, 5,* 373-390. https://doi.org/10.1007/s13164-014-0193-4

Fernández, J. (2019) *Memory: a self-referential account.* Oxford University Press. https://doi.org/10.1093/oso/9780190073008.001.0001

Grice, H. P. (1941). Personal identity. *Mind, 50,* 330-350. https://doi.org/10.1093/mind/L.200.330

Hamilton, A. (2007). Memory and self-consciousness: immunity to error through misidentification. *Synthese, 171,* 409-471. https://doi.org/10.1007/s11229-008-9318-6

Higginbotham, J. (2010) On words and thoughts about oneself. In F. Recanati, I. Stojanovic & N. Villaneuva (Eds.)*, Context-dependence, perspective, and relativity* (pp. 253-282)*.* De Gruyter Mouton.

Horgan, T. & Kriegel, U. (2007). Phenomenal epistemology: what is consciousness that we may know it so well?*. Philosophical Issues, 17,* 123-144. https://doi.org/10.1111/j.1533-6077.2007.00126.x

Ismael, J. (2012). Immunity to error as an artefact of transition between representational media. In S. Prosser & F. Recanati (Eds.), *Immunity to error through misidentification: new essays.* Cambridge University Press.

Levine, J. (2018). Conscious awareness and (self-) representation. In *Quality and content: essays on consciousness, representation, and modality. Oxford University Press.* https://doi.org/10.1093/oso/9780198800088.003.0007

McCarroll, C. (2018). *Remembering from the outside: personal memory and the perspectival mind.* Oxford University Press. https://doi.org/10.1093/oso/9780190674267.003.0007

Merlo, G. (2017). Three questions about immunity to error through misidentification, *Erkenntnis, 82,* 603-623. https://doi.org/10.1007/s10670-016-9834-6

Michaelian, K. (2020). Episodic memory is not immune to error through misidentification: against Fernández. *Synthese.* https://doi.org/10.1007/s11229-020-02652-w

Millar, A. (1991). *Reasons and experiences.* Clarendon. https://doi.org/10.1093/acprof:oso/9780198242703.001.0001

Olson, E. T. (2019). Personal identity. In E. N. Zalta (Ed.). *The Stanford Encyclopedia of Philosophy* (Spring 2021 Edition), https://plato.stanford.edu/archives/spr2021/entries/identity-personal/

Parfit, D. (1984) *Reasons and persons.* Clarendon.

Estud.filos n.º 64. Julio-diciembre de 2021 | pp. 189-200 | Universidad de Antioquia | ISSN 0121-3628 | ISSN-e 2256-358X

199

https://doi.org/10.17533/udea.ef.n64a10

Pasupathi, M. & Wainryb, C. (2018). Remembering good and bad times together: functions of collaborative remembering. In M. Meade, C. Harris, P. Van Bergen, J. Sutton & A. Barnier (Eds.), *Collaborative remembering: theories, research, and applications.* Oxford University Press. https://doi.org/10.1093/oso/9780198737865.003.0015

Pillemer, D. B., Steiner, K. L., Kuwabara, K. J., Thomsen, D. K. & Svob, C. (2015). Vicarious memories. *Consciousness and cognition, 36,* 233-245. https://doi.org/10.1016/j.concog.2015.06.010

Prosser, (2012). Sources of immunity to error through misidentification. In S. Prosser and F. Recanati (Eds.), *Immunity to error through misidentification: new essays.* Cambridge University Press. https://doi.org/10.1017/CBO9781139043274

Pryor, J. (1999). Immunity to error through misidentification. *Philosophical Topics, 26,* 271-304. https://doi.org/10.5840/philtopics1999261/246

Recanati, F. (2007). *Perspectival thought: a plea for (moderate) relativism.* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199230532.001.0001

Reese, E. & Brown, N. (2000). Reminiscing and recounting in the preschool years. *Applied Cognitive Psychology, 14.* 1-17. https://doi.org/10.1002/(SICI)1099-0720(200001)14:1<1::AID-ACP625>3.0.CO;2-G

Ross, L. & Ward, A. (1996). Naïve realism: implications for social conflict and misunderstanding. In T. Brown, E. Reed & E. Turiel (Eds.), *Values and Knowledge* (pp. 103-35). Lawrence Erlbaum.

Schacter, D. L., & Addis, D. R. (2007). The cognitive neuroscience of constructive memory: remembering the past and imagining the future. *Philosophical Transactions of the Royal Society B, 262,* 773–786. https://doi.org/10.1098/rstb.2007.2087

Sheen, M., Kemp, S. & Rubin, D. (2001). Twins dispute memory ownership: a new false memory phenomenon. *Memory & Cognition, 29*(6), 779-788. https://doi.org/10.3758/BF03196407

Shoemaker, S. (1959). Personal identity and memory. *The Journal of Philosophy, 56*(22), 868-882. https://doi.org/10.2307/2022317

Shoemaker, S. (1968). Self-reference and self-awareness. *Journal of Philosophy, 65,* 555-567. https://doi.org/10.2307/2024121

Shoemaker, S. (1970). Persons and their pasts. *American Philosophical Quarterly, 7,* 269-285.

Soteriou, M. (2000). The particularity of visual perception. *European Journal of Philosophy, 8,* 173-189. https://doi.org/10.1111/1468-0378.00107

Tye, M. (2009). *Consciousness revisited: materialism without phenomenal concepts.* MIT Press. https://doi.org/10.7551/mitpress/9780262012737.001.0001

Wittgenstein, L. (1958). *The blue and brown books.* Blackwell.

The failures of functionalism (for memory)*

Sarah Robins

University of Kansas, Kansas, United States
E-mail: skrobins@ku.edu

Abstract: In *Memory: A Self-Referential Account,* Fernández offers a functionalist account of the metaphysics of memory, which is portrayed as presenting significant advantages over causal and narrative theories of memory. In this paper, I present a series of challenges for Fernández's functionalism. There are issues with both the particulars of the account and the use of functionalism more generally. First, in characterizing the mnemonic role of episodic remembering, Fernández fails to make clear how the mental image type that plays this role should be identified. Second, I argue that a functionalist approach, which appeals to the overall structure of the memory system and tendencies of mental state types, is ill-suited to the metaphysical question about episodic remembering that is of interest to the causal and narrative theorists with which Fernandez engages. Fernández's self-referential account of memory has many other virtues, but functionalism is a poor fit for episodic remembering.

Key words: memory, functionalism, episodic memory, mental image

Cómo citar este artículo

Robins, S. (2021). The failures of functionalism (for memory). *Estudios de Filosofía, 64,* 201-222. https://doi.org/10.17533/udea.ef.n64a11

ARTÍCULO
DE REFLEXIÓN

Los fallos del funcionalismo
(para la memoria)

Resumen: En *Memory: A Self-Referential Account,* Fernández ofrece una explicación funcionalista de la metafísica de la memoria, que se dice tener ventajas significativas sobre las teorías causales y narrativas de la memoria. En este artículo presento una serie de desafíos para el funcionalismo de Fernández. Hay problemas tanto con los detalles del relato como con el uso del funcionalismo en general. En primer lugar, al caracterizar el papel mnemónico del recuerdo episódico, Fernández no aclara cómo debe identificarse el tipo de imagen mental que desempeña este papel. En segundo lugar, sostengo que un enfoque funcionalista, que apela a la estructura general del sistema de memoria y las tendencias de los tipos de estados mentales, no se adapta a la pregunta metafísica sobre el recuerdo episódico que es de interés para los teóricos causales y narrativos con los que Fernández discute. La descripción autorreferencial de la memoria de Fernández tiene muchas otras virtudes, pero el funcionalismo no encaja bien con el recuerdo episódico.

Palabras claves: memoria, funcionalismo, memoria episódica, imagen mental

Sarah Robins

is an Associate Professor of Philosophy at the University of Kansas and an affiliate of the Cognitive and Brain Sciences Ph.D. program in KU's Psychology Department. Her research is focused primarily on memory, with an emphasis on memory traces and the ways they are invoked and investigated in philosophy, psychology, and neuroscience.

ORCID: 0000-0003-2439-4633

Check for updates

1. Introduction

Fernández's *Memory: A Self-Referential Account* is ambitious, tackling questions about the metaphysics, intentionality, phenomenology, and epistemology of memory simultaneously. The result is a bold and broad theory of remembering, full of arguments and ideas that will no doubt influence the philosophy of memory for years to come. Fernández has built the account so that its central components are relatively free-standing —he claims that one can, for example, accept his proposed view of memory content while rejecting the account of its metaphysics, or vice versa. This design allows me to focus my critique on the metaphysical account of memory that Fernández proposes, specifically the functionalist view he sets out in chapter two.

Fernández's functionalism characterizes experiences of episodic remembering in terms of the mnemonic role played by their mental images. When a mental image fulfills this mnemonic role, the subject is remembering. In this paper, I present a number of challenges to this functionalist proposal. First, I argue that there are a number of difficulties with the particulars of Fernández's functionalism, which cause problems for the mnemonic role as he's characterized it. Second, I argue that there are broader problems for the attempt to apply a functionalist framework to episodic remembering. Functionalism offers a characterization of mental state types and allows for assessments of remembering across individuals, in terms of whether or not they possess that mental state type. While on the surface they may look similar, these concerns are importantly distinct from those at issue for causal and narrative theorists, who are proposing conditions on token states of remembering that will make possible assessments of remembering within individuals. Ultimately, I conclude, functionalism fails to fit the explanatory demands of episodic remembering.

2. Fernández's functionalism

Fernández's account of memory is concentrated on the intersection between memory for facts and memory for perceptual experiences. Focusing on memory for facts, or propositional memory, is fairly common amongst philosophers offering accounts of memory (e.g., Bernecker, 2010). Fernández's (2019) approach is unique because he narrows the scope of his account to a subset of propositional memory: "memory for facts involving objects perceivable through sensory modalities" (p. 5). In the book, he focuses specifically on vision and visible objects —a memory of the fact that the keys were left on the counter, or a memory of the fact that Mary attended the party.

Fernández then narrows his focus further. His interest is in how facts about perceptible objects are remembered *episodically.* This involves appeal to the familiar distinction between episodic and semantic memory, but Fernández puts it to a particular use. The distinction between episodic and semantic memory is often cast as

one between memory for experiences and memory for facts, respectively. Fernández is already calling attention to ways that distinction can be blurred with his interest in memory for facts based in perceptual experience. He uses the distinction between episodic semantic to characterize forms of *remembering,* the activity of entertaining a memory. Episodic remembering involves experiences; semantic remembering involves beliefs. Even for memories of facts based in perceptual experience, both forms of remembering are possible. My remembering, now, that they keys are on the counter could be the activation of a belief that I formed when I saw the keys on the counter earlier in the day. My remembering now could also be experiential: having just noticed the keys are not with me, I may visualize the rooms of my house, use my attention to scan various objects lying around, and in so doing, locate my keys on the counter. This latter form is remembering episodically.

Episodic remembering, for Fernández (2019), is experiential: "when you remember [episodically], you are having an experience; an experience that typically originates in a past perceptual experience of the fact" (p. 6). Episodic remembering, like other experiential mental states, has phenomenal and intentional features. Fernández discusses these features later in the book. First, he addresses a complementary metaphysical question about episodic memory: What are the conditions under which an experience qualifies as an episodic memory?[1] The candidate experiences are mental images —the image of keys on the counter or the image of Mary amongst the party attendees. But this alone does not answer the question. Lots of mental states involve mental imagery, not just remembering. What identifies the mental images of remembering uniquely?

Answering this question is difficult, in part, because of the diachronic nature of memory. Episodic remembering is a mental activity *now* but which is about and, in some sense, due to a previous experience. The metaphysical conditions on episodic remembering must therefore include consideration of the past and the present. In other words, the metaphysics of memory involves saying, of a current mental image, something about where it came from *and* what is being done with it now. Fernández's discussion begins with discussion of two metaphysical views —the causal theory of memory and the narrative theory of memory. Each of these alternative views emphasizes one aspect of memory's diachronic nature. Causal accounts are focused on the connection to the past event; narrative accounts are focused on the role of the state in one's present mental life. Fernández argues that these approaches, while distinct, have parallel flaws: each over-emphasizes its selected dimension and neglects the importance of the other. Fernández thus offers functionalism, a theoretical approach that appeals to

---

1   Fernández (2019) characterizes his view as a hybrid approach that pursues metaphysical and intentional features of memory to be equally fundamental. He of course views the account he develops of the two approaches to be complementary, but acknowledges that the two are meant to function independently (p. 24).

both causes *and* effects, as an account better suited to accommodating both aspects of the diachronic nature of episodic remembering. He then demonstrates its superiority by showing how functionalism can remedy the deficiencies of each alternative, better addressing the cases that were problematic for both causal and narrative theories.

In what follows, I provide a brief summary of Fernández's characterization of each of the alternative theories and the cases that present problems for each. These cases play a critical role in Fernández's own account, as functionalism's key strength is its alleged ability to offer better responses to them. In discussing these cases, I use the same examples as Fernández does in his original presentation, so as to best ensure that the relevant features are preserved.

2.1 Causal theories of memory

Causal theories of memory privilege the connection between the episodic remembering experience and the past event (e.g., Bernecker, 2010; Debus, 2010; Martin & Deutscher, 1966). More specifically, they require a causal connection between the event and the subsequent experience. Versions of the causal theory differ over how they identify the particular kind of causal connection that is required. What all versions share, Fernández argues, is too rigid of a focus on the past event and its influence on remembering. This leads the causal theory to omit some cases of remembering where the connection to the past is weaker, and to include cases that have this connection even when they lack features that are intuitively essential for remembering. Cases of these sorts Fernández labels *embellishment* and *epistemic irrelevance,* respectively.

Embellishment cases are ones where the content involved in the experience of remembering goes beyond what was available in the previous experience. Fernández uses the example of a person hunting with their father who, on one occasion, sees their father shoot a white rabbit. Later, the person visualizes their father having shot a black rabbit. Despite the change in content, the case should still count as a case of remembering, albeit one that involves *misremembering.*[2]

Epistemic irrelevance cases involve a person generating a mental image that derives from a past experience, but that they do not recognize as such. Fernández uses the example of a person who is painting, and draws a bird sitting atop a house. Unbeknownst to the painter, this scene is one they saw as a child. Since the painter is unaware of this connection to his past, he does not engage with the mental image (or the painting) in ways characteristic of remembering. He does not, for example,

---

2   Here it is important to note that this way of classifying states makes sense on Fernández's view because remembering is not a factive state. For Fernández, it is possible to be remembering and do so incorrectly.

believe that the scene depicted is one that he previously experienced. Since the image does have a connection to his past experience, causal theories admit such cases. But such cases lack several important features —the experiential feeling of remembering, an associated belief that the image derives from past experience, etc.— which should lead us to preclude such epistemic irrelevance cases, Fernández argues.

2.2 Narrative theories of memory

Narrative theories of memory, in contrast, focus on how experiences of episodic remembering are integrated into a person's sense of their life and its temporally outstretched narrative (e.g., Brockmeier, 2015; Goldie, 2012; Schectman, 1994). The emphasis is on the role the mental image plays in the person's sense of their self and their life, not on the connection to the past. The view is thus well-suited to accommodate the embellishment cases that causal theories neglected, while also requiring epistemic relevance that would block cases like the painter. The view encounters troublesome cases of its own, however, because of its overemphasis on integration and because of its failure to require a connection to the past. Narrative theories are thus susceptible to cases of *isolation* and *confabulation.*

Isolation cases are ones where a person has a vivid mental image of a past experience, that they recognize as a memory and believe to be a depiction of something that previously happened to them, but where the person is simultaneously unable to situate the experience at any particular point in the past. Fernández's example involves falling into a pool at some point during childhood. The person in the example vividly recalls the experience, but cannot remember when this occurred, where the pool was, who was there, etc. The person's failure to integrate this experience with his broader life narrative precludes it from the class of rememberings for supporters the narrative theory. Even if such memories are isolated, Fernández argues, they should still be included.

 Confabulation cases are the inverse of epistemic irrelevance cases. They occur when a person has a mental image that is treated as a memory, and well-integrated into the person's understanding of their prior experience and life narrative —despite the fact that the mental image does not derive from a past experience. Confabulation often occurs in clinical cases of psychiatric disorder, and indeed, Fernández's example of such a case involves a patient with Korsakoff's syndrome who, despite having amnesia and being hospitalized, claims to have had a conversation on the train while traveling over the weekend. Such a case meets the narrative theory's requirements for integration, but lacks any connection to the past experience. Narrative theories thus allow such cases to count as instances of remembering, but they should not.

2.3 Memory functionalism

Fernández's presentation of causal and narrative theories of memory, and their respective limitations, provides an elegant setup for generating a new account of the metaphysics of memory. Each account has strengths and weaknesses, which complement one another. The aim is for a middle-ground account that can maintain both sets of strengths while avoiding the accompanying weaknesses. An adequate account of what is required for a mental image to qualify as an experience of remembering episodically must include both a connection to the past and integration into one's present life, while at the same time stopping short of demanding complete fealty to the past or wholesale enmeshing into one's life narrative.

Functionalism, Fernández argues, can provide such an account. Functionalist analyses are common amongst accounts of mental states, providing an important precedent for extending this framework to memory. Functionalism involves a characterization of the mental state in question in terms of its functional role, generally understood as a specification of the state's typical causes and typical effects. By appealing to both causes and effects, functionalism meets the requirement of connecting to both the past and the present. Additionally, the appeal to *typical* causes and effects gives the account flexibility, helping to ensure that neither connection is interpreted rigidly enough to cause the problems that were shown above for causal and narrative accounts.

Episodically remembering a fact about one's past perceptual experience is, for Fernández, a matter of having a mental image that plays the right functional role. He characterizes the requisite mnemonic role as follows:

> *S* remembers that *p* just in case *S* has some mental image *i* such that *i* tends to cause in *S* a disposition to believe both that *p* and that *S* experienced that *p*, and *i* tends to be caused in *S* by having experienced that *p* (Fernández, 2019, p. 49).[3]

The account is centered upon the mnemonic role of a mental image, *i.* The account includes a connection to the past and a connection to the present. To play the requisite role, *i* has to have certain causes *and* certain effects. These relations to the past and present are, however, framed as tendencies, allowing minor aberrations in a way that helps in handling trickier cases like embellishment and isolation.

Fernández argues that these features allow functionalism to retain the successful qualities of both causal and narrative accounts, while also addressing the cases that caused trouble for these alternatives. Functionalism can, he claims, accommodate embellishment and isolation, while excluding epistemic irrelevance and confabulation. The required connection to the past prevents confabulation; the required connection to the present prevents epistemic irrelevance. Construing both required connections

---

3   Fernández (2018) offers an initial account of functionalism.

as tendencies leaves enough wiggle room to allow embellishment and isolation into the account of remembering. Or so Fernández argues.

Before turning to an evaluation of Fernández's functionalist proposal, it is worth pausing to reflect on its significance. Fernández's functionalism brings a novel perspective to the metaphysics of memory, re-energizing a debate where the number of viable positions has long been limited. By promoting functionalism, Fernández also encourages philosophers interested in the metaphysics of memory to engage with material available in the metaphysics of mind more broadly. Functionalism also offers a few particular advantages. First, by characterizing the mental state of remembering in terms of its mnemonic role, Fernández is able to sidestep murky debates over the mental content of episodic memory. Second, Fernández puts the requirement of connection to the present on equal footing with the requirement of connection to the past. This is rare amongst accounts of memory, which are understandably focused on the past as the source from which remembering derives. In so doing, he helps to highlight the importance of a memory's relevance in a person's cognitive and epistemic activities at the time of remembering that past-directed accounts often overlook.[4]

Having laid out Fernández's functionalism and its alleged advantages over alternative accounts, I now turn to critiquing the view. My critique comes in two forms, addressed in the following two sections. First, I focus on the particulars of Fernández's account —whether the mnemonic role for memory images, as formulated, can do the work he intends for it. Second, I introduce a more general set of concerns about the use of functionalism to provide a metaphysics of memory. I argue that the structure and strengths of functionalism are ill-suited to the concerns at issue for the causal and narrative theorists with which Fernandez is engaged.

3. Mental images and the mnemonic role

When looking into the details of Fernández's view, my concerns are focused on the mnemonic role he sets out for mental images —specifically, how to individuate the mental state type *i* that's meant to play this role. Fernández's defense of functionalism is brief: he offers a characterization of the mnemonic role, quoted above, and some remarks about how this approach can improve upon the deficiencies of the causal and narrative approaches. His presentation does not, however, involve walking through any of the examples in detail, explicating how each variable in the mnemonic role is filled and how its tendencies are evaluated. This is understandable, given how many aspects of memory Fernández addresses in this book-length account. Taking the time

---

4   Although see Debus (2010) as an example of a causal theorist who has incorporated concerns about relevance.

to go through some examples in detail, as I do below, raises questions that need to be addressed before the account can deliver on its promised advantages.

Let's start with a straightforward case, derived from one of Fernández's own examples: episodically remembering the fact that Mary was at the party last month. In this case, *S* has an experience of a mental image —presumably, an image of the party, with Mary amongst the attendees. What is required for this experience to qualify as one of episodic remembering? To answer the question, Fernández's functionalism asks us to consider *i*, the mental image of Mary at the party, and its role in *S*'s mental life. What tends to cause this image and what effects does it tend to produce?

In order to evaluate *i*'s tendencies, we need to situate this token within its broader mental state type. Fernández does not tell us how this is to be done, but his view contains material that provides some suggestions. It seems clear that the mental state type in question will be *subject-relative*. The mnemonic role is identified through *i*'s tendencies *in S*. This makes sense. These are images of experiences, from a visual perspective occupied by *S*, not anyone else. People differ not only in the perspectives they have on any particular fact in the world, but more broadly in terms of when and where they live, and so, which mental images they have and what they tend to cause and effect. I have a mental image of my first day of college. My son, who was not alive at the time, does not. The mental image types that operate as *i* for me will be different than the ones that operate for my son, which in turn will differ from yours and everyone else's.

Even once the relevant mental image type has been restricted to a given subject *S*, there are still multiple ways it could be characterized. It could be the set of all occurrences of *i* —the collection of all the times the mental image has been tokened in *S*'s experiences. Take the case of remembering that Mary was at the party. Suppose the first time *S* tokened this mental image it was as part of visualizing the party scene to determine whether Mary was in attendance. Subsequently, however, this is the mental image that comes to mind whenever Mary's name comes up in conversation or whenever *S* wonders how Mary is doing. This party may be the last time *S* saw Mary, and as more time passes *S* begins to wonder what has happened to Mary and what has gone wrong with their friendship. The mental image type *i* now has several tokened instances, making it easier to evaluate its tendencies. The problem, at least for *i*'s mnemonic role, is that *i* doesn't have the tendencies required for episodic remembering. In the first instance, when *S* visualized the party to scan the image for Mary, this seemed like an instance of remembering. But *i* does not tend to produce the belief that *S* experienced Mary being at the party. It had this effect once, but now it is more likely to be involved in the production of other mental states —worries that Mary is upset with S, or a belief that Mary is avoiding *S*. To put the worry more generally: as the evaluation shifts to *i* as a mental state type, any tendencies that would have supported its mnemonic role are weakened.

Estud.filos n.º 64. Julio-diciembre de 2021 | pp. 201-222 | Universidad de Antioquia | ISSN 0121-3628 | ISSN-e 2256-358X | 209

https://doi.org/10.17533/udea.ef.n64a11

Given this problem for this formulation of *i*, we could try another approach — instead of all of the ways *S has* tokened this mental image of Mary at the party, *i* might be all of the ways that *S could* do so. Fernández speaks in this way in some of the cases he considers, where the relevant roles are described as possible situations or what would have happened had S seen a particular thing other than what was actually seen. This modification provides the account with more flexibility, but it is unclear how to constrain the realm of possible cases of *S* visualizing *i*. Are we to assume that *S*, and her place in space-time, are fixed? Are we holding constant all of her activities, experiences, and beliefs up to this moment? If too many features change, then *i* becomes difficult to evaluate —*S* could fail to attend the party, or not know Mary, etc. The more of these factors that are open, the more flexibility the account gains.

What is needed is a way of explaining which possible tokenings of *i* are relevant, one that offers more than simply selecting a few cases we would like to consider for a particular evaluation. Fernández does not provide this. Regardless of which way it is specified, the flexibility it provides seems at odds with securing *i*'s mnemonic role. Once we pause to consider all of the possible ways a mental image of Mary at the party could be put to use by *S*, we become aware of just how many effects (and causes) this state could have. The wider the set of options available, the more the tendency toward the generation of any particular belief as an effect becomes increasingly small.

Neither of these candidates for *i* —the set of all of *S*'s experiences of visualizing *i*, the set of all ways *i* could be visualized by *S*— fit into the mnemonic role as Fernández has sketched it. Other alternatives may come from Fernández's discussion of how functionalism can better handle the cases that posed problems for causal and narrative accounts. Let's start, as Fernández does, with embellishment. Embellishment cases are ones where the content involved in the act of remembering goes beyond (i.e., provides more detail than or alternative detail to) what actually occurred. Fernández's example involves an *S* who, in the past, saw their father shoot a white rabbit. *S*'s mental image now, however, is of their father shooting a black rabbit. Fernández believes such cases should be counted as instances of remembering, albeit ones that involve misremembering. A functionalist account can accommodate such cases, Fernández argues, thanks to its reliance on tendencies rather than what actually occurred. We should thus be able to see how the mental image *i* is understood in this embellishment case so that it succeeds in filling the mnemonic role.

Recall Fernández's account of the mnemonic role for *i* involves tendencies in two directions: what tends to cause *i* and what effects *i* tends to produce. When considering the effects *i* tends to produce, Fernández appears to have a very specific image in mind —*S*'s an image of their father shooting a black rabbit. It fulfills this half of the requisite mnemonic role because

> my mental image tends to cause in me the belief that I once saw a black rabbit being shot by my father, and it tends to cause in me the belief that my past perceptual experience was veridical; that the shooting did obtain in the past (Fernández, 2019, p. 51).

Producing the right effects requires that *i* involve the kind of details that would support belief about this particular past event as an experience that S has had.

When we turn to the second set of tendencies to evaluate for *i* —what tends to cause this mental image to come about— the *i* under consideration appears to change. Here is what Fernández says:

> My mental image is the type of image that tends to be produced in me by past perceptual experiences of black rabbits being shot. To be sure, on this particular occasion, my mental image was not actually caused by a perceptual experience of a black rabbit being shot since, in the past, I did not have such an experience. Nevertheless, the fact remains that my faculties of perception and memory are related in such a way that perceptual experiences of black rabbits do produce in me the type of mental image that I am currently having. Had I seen, in other words, a black rabbit being shot in the past, this is the type of mental image I would be having now (Fernández, 2019, p. 51).

To generate a case where *i* has the tendencies necessary for this half of its mnemonic role, Fernández considers *i* as a generic image type: mental images of the type S tends to have when S sees black rabbits being shot. In the example, we are not given enough detail about S to know how many times this has happened for *S*, whether it is a common perceptual experience for *S* or indeed whether *S* has ever had this experience. Suppose that S has seen black rabbits being shot several times in the past. We might first wonder what type of mental image this would be: presumably, it is some generic image amalgamated from each of these experiences, with the details about the particular size of the rabbit, location of its wound, nearby ground cover, weather, etc. somehow smoothed out. However this goes, the *i* that emerges from the collection of all the times *S* has seen black rabbits will differ from the *i* that produces, in *S*, the belief that S had the past experience of their father shooting a black rabbit. Producing the belief required perceptual details about a particular experience; the image type associated with seeing black rabbits shot would seem to lack both these details and a connection to a particular experience.

The above quotation from Fernández also suggests a different reading of *i* and what tends to cause it. This alternative doesn't rely on S having had multiple experiences of seeing black rabbits shot in the past. Instead, it appeals to the general workings of and connections between *S*'s perceptual and memory systems to establish what would have happened if *S* had seen a black rabbit being shot.

Estud.filos n.º 64. Julio-diciembre de 2021 | pp. 201-222 | Universidad de Antioquia | ISSN 0121-3628 | ISSN-e 2256-358X | 211

https://doi.org/10.17533/udea.ef.n64a11

Fernández does not elaborate on how this evaluation of perception and memory should go, but we can make a straightforward guess. *S* has a well-functioning visual system, so had she seen a black rabbit being shot, she would be likely to see it correctly. Further, *S*'s memory system is well connected to her visual system; her memory tends to store images of what she saw in this past. So had she seen a black rabbit, she would have been likely not only to see it correctly, but to retain a mental image of that experience. These judgments seem fine enough, and we can stipulate that they are true of the *S* in the example.

What remains unclear, however, is how consideration of this counterfactual gives us an evaluation of *i*'s tendencies that yields the desired conclusion. To secure *i*'s mnemonic role, it must be the case that *i* tends to be produced by *S*'s experience of *p* (here *p* = *S*'s father shooting a black rabbit). In actuality, *i* was produced by *S*'s father shooting a white rabbit, but evaluation of the counterfactual shows another way of producing *i*. In short, *i* came about in one way, but it could have come about another way. Considering these two cases together, what does that tell us about *i*'s tendencies? It is unclear to me how we are supposed to arrive at any conclusion about its tendencies, much less the conclusion that *i* tends to be produced by seeing a black rabbit being shot. And either way, this interpretation of *i* does not help with the problem previously identified about reconciling the senses of *i* used to determine its causes and its effects.

To put the point from embellishment cases succinctly: evaluating *i*'s mnemonic role involves assessing two tendencies —what tends to cause *i* and what *i* tends to cause. The understanding of *i* used to establish one side of its mnemonic role looks different than the understanding of *i* used to establish the other. In order for *i* to produce the right beliefs in *S*, *i* has to be very specific. This is the only way for *i* to tend to cause beliefs that a particular perceptual experience occurred. But in order for *i* to be caused by the right sort of experience, it needs to be generic. There does not appear to be a stable conception of *i* that can play both roles.

The difficulty in providing a stable conception of *i* across both sides of its mnemonic role creates further problems for Fernández's treatment of confabulation cases —and the ability of his functionalism to keep cases of embellishment and confabulation separate.

Confabulation cases are ones where a mental image is integrated into a person's thoughts and actions *as* if it were a memory, but where the image has no connection to the person's past experience. Fernández uses the example of a person with Korsakoff's syndrome, a memory disorder induced by extreme thiamine-deficiency, often as a result of heavy alcohol use. In the example, the Korsakoff's patient is hospitalized, and has been for some time, but when asked about their activities on the prior day, the patient generates an elaborate account of going on a trip and having an extended conversation on the train. The patient has a mental image *i* of

conversing with a fellow passenger on a train while taking a trip. This image fulfills one half of the requisite mnemonic role: the image tends to produce beliefs in the patient that they were on a train over the weekend, having a conversation with a fellow passenger. This tendency led to confabulation cases being accepted by narrative theories of memory. Fernández argues, as many memory theorists do, that cases of confabulation should be excluded from an account of remembering. In order to exclude such a case, the operative question is whether "the patient's mental image of the conversation in the train is of a kind which does not tend to be produced, in that patient, by experiences of such conversations" (Fernández, 2019, p. 52).

Fernández believes that the answer to this question is no. The Korsakoff's patient has amnesia. As such, they lack the kinds of tendencies required. For them, mental images active during perception do not tend to be stored and converted into memories. Without such a connection, *i* fails to fill the mnemonic role, and cases of confabulation fail to count as instances of remembering.

At first glance, this judgment on confabulation cases looks consistent with Fernández's treatment of embellishment cases, as discussed above. That is, Fernández is using an assessment of *S*'s faculties of perception and memory to determine whether *i* tends to be caused in the right way. Embellishment and confabulation both involve errors: the mental images offer an incorrect depiction of *S*'s past experience in both cases. Despite this similarity, embellishment cases count as remembering while confabulation cases do not because of the broader tendencies that exist in the cognitive systems from which they are generated. Embellishment cases count as (mis)remembering because they emerge from functional perceptual and memory systems; confabulation cases fail to count because they emerge from dysfunctional perceptual and memory systems.

This approach to confabulation is difficult to reconcile with the literature on Korsakoff's and other forms of amnesia. Korsakoff's patients, almost universally, have *anterograde* amnesia −the onset of the disorder is marked by the inability to form and store new memories (Fama, Patel & Sullivan, 2012). There are significant differences amongst persons with Korsakoff's, however, on the extent of additional memory damage. There is a tendency for at least some retrograde amnesia (Kopelman, Thomson, Guerrini & Marshall, 2009), but the extent of a person's inability to remember past events differs across individuals. Fernandez responds to all persons with Korsakoff's in the same way, dismissing them from possible remembering because of their lack of properly functioning perceptual and memory systems. Assessment of persons with Korsakoff's, however, and the determination of how much of their memory systems are or are not functioning proceeds individually. Persons who are diagnosed with Korsakoff's often present with confabulations about recent events. Spiegel and Lim (2011) describe an individual who reported that he had just arrived on a flight from out of state (when in fact he was in the

Estud.filos n.º 64. Julio-diciembre de 2021 | pp. 201-222 | Universidad de Antioquia | ISSN 0121-3628 | ISSN-e 2256-358X   213

https://doi.org/10.17533/udea.ef.n64a11

emergency room at the hospital, as a result of falling from a balcony) and that Barack Obama, the US President at the time, had died. From initial confabulations such as these, doctors go on to administer additional tests, in order to determine whether or not the patient's retrograde memory systems are also damaged, and if so, how severely.

It is thus consistent with a Korsakoff's diagnosis for a person to retain memories from the past, maintaining some functional connections between perception and memory. Acknowledging this complicates the assessment of whether any particular mental image produced by a person with Korsakoff's is an instance of confabulation. In the case Fernández describes, the patient has a mental image of a recent event, one that would have occurred after the development of the disorder and so would not have the tendency to be the result of a past perceptual experience. But suppose this patient had (prior to developing Korsakoff's) previously traveled a lot, often doing so by train, and on some occasions, talked with fellow passengers. In this case, the patient may very well have a mental image i that has the tendencies needed in order to play the mnemonic role. That is, the mental image of talking to other travelers on a train is one that tends to be caused in S by having had this experience. In this particular instance, i was not caused in this way —but the appeal to tendencies is meant to provide a way to step back from what actually happened when making our evaluation. For many confabulation cases, then, extending the evaluation of i and its mnemonic role that was used to assess embellishment cases, will yield the conclusion that confabulations are instances of remembering.

Across all of these cases, what matters for determining whether it counts as remembering is how the mental image i is characterized. To secure its role in acquainting the rememberer with a fact about past perception, there is pressure for i to be highly specific. To evaluate its broader tendencies, allowing for hiccups in the standard process of remembering, there is pressure in the opposite direction, for i to be highly general. Fernández does not offer much in way of elaboration on i. What he does provide pulls in both directions, as illustrated above. Is there a way for i to retain its connection to a particular past experience while still allowing a way for its tendencies to be evaluated? Moreover, can this formulation of i allow in cases of embellishment while keeping out cases of confabulation?

Fernández, no doubt, has more to say about these cases and the resources within functionalism for responding to the challenges presented here. It is possible that there are ways to modify the mnemonic role, or the understanding of i to address these issues. Whether functionalism can withstand these challenges and make good on its promises as a metaphysics of memory is thus yet to be determined. For now, I will set these concerns about the details of functionalism aside and turn to a broader critique of the use of a functionalist framework to account for episodic remembering.

4. Functionalism's failure of fit

In the previous section, I challenged Fernández's account of the mnemonic role for episodic remembering, arguing that his characterization lacked the detail required to sort through particular cases. Viewed from another angle, this alleged flaw is actually a feature, one that is characteristic of functionalist accounts of mental states. Fernández's functional characterization of episodic remembering is highly similar to functionalist characterizations of other mental states, like belief (e.g., Leitgeb, 2017) and desire (e.g., Alvarez, 2017). These general sketches of the role played by belief and desire do the work asked of them; theorists recognize the full specification of any particular belief or desire will be longer and more complicated.[5] In other words, I have been criticizing Fernández's functionalist view for being functionalist. This could indicate that my objections to the view are off base. Instead, I think this observation opens the door to a distinct way of evaluating Fernández's project, one that highlights the tension between the aims and interests of functionalism and those that have standardly been the focus for other philosophers of memory, particularly those endorsing causal and narrative views.

As an account of the metaphysics of remembering, Fernández's functionalism is an answer to the following question: *what are the conditions under which an experience qualifies as an episodic memory?* On the surface, this looks to be the same question asked (and answered) by causal and narrative theorists. In the details, however, they are importantly different. Each asks the question in a way that invokes a distinct contrast class, placing divergent constraints on what counts as an adequate answer.

When Fernández's asks about the conditions on episodic remembering, he is looking for a way to distinguish episodic memory —as a mental state type— from other mental state types. As he says at the outset of the chapter, his aim is "to determine what it is to remember something, as opposed to imagining it, perceiving it, or introspecting it" (Fernández, 2019, p. 32). Causal and narrative theorists, in contrast, are interested in exploring these conditions as a way of distinguishing remembering from borderline and/ or degenerate cases. In other words, they are asking: amongst the set of mental states that are candidates for remembering, which ones are successful, genuine, or real? Some theorists approach this question by treating the sought-after category of successful remembering as factive. The aim is then to identify all and only the cases of apparent remembering in which the past event is accurately represented. The criterion need not be accuracy, though, as illustrated by narrative accounts that appeal to integration into one's life story as critical for genuine memory. Regardless of how successful cases are defined, the conditions used to establish them are ones that allow sorting between token mental states.

There are of course broad similarities between Fernández's question and the question asked by causal and narrative theorists. They're all discussing the conditions on episodic remembering, and moreover, they are all engaged with this question as a metaphysical project. The question is about the conditions that must obtain for remembering, not one of how we could know whether they are operative in a particular case or not. The accounts differ in terms of the scope at which the question is pitched —are the conditions placed on a mental state type or on particular token states? This difference in scope has consequences for how the question is explored and answered; putting the two into direct conversation, as Fernández's does here, leads to confusion and to problems.

The tension is apparent in Fernández's account. Although he frames the chapter's question as one of comparing remembering to other mental state types, the account he provides does not answer that question. Nothing in Chapter 2 addresses how memory differs from perception, imagination, or other imagery-involving mental states. Instead, his account develops out of engagement with causal and narrative approaches, where he uses functionalism to address borderline cases of embellishment, isolation, confabulation, and epistemic irrelevance. Despite focusing on these cases, which for causal and narrative theorists require sorting amongst tokens to determine which cases are successful or genuine, Fernández's maintains his approach based in remembering as a mental state type. As such, the answers he gives are ill-suited to the questions these cases raise, at least as they are standardly understood by causal and narrative theorists. Given this mismatch, Fernandez's account fails to answer either form of the question about the conditions on episodic remembering.

To illustrate this, I will start by restating Fernández's characterization of the mnemonic role of episodic remembering:

> S remembers that *p* just in case *S* has some mental image *i* such that *i* tends to cause in *S* a disposition to believe both that *p* and that *S* experienced that *p,* and *i* tends to be caused in *S* by having experienced that *p* (Fernández, 2019, p. 49).[6]

Engaging with causal and narrative views, as he does in this chapter, requires Fernández's to put this mnemonic role to work evaluating particular cases. Fernández's considers four: embellishment, isolation, confabulation, and epistemic irrelevance. As discussed in Section 2, the goal is to develop an account that includes the first two (embellishment and isolation), while excluding the latter two (confabulation and epistemic irrelevance). When Fernandez introduces these cases, he includes specific, detailed examples —a hunting trip where one's father shoots a rabbit (embellishment), falling into a swimming pool as a child (isolation). Generating cases in this way is done to mirror the features involved in cases generated by causal and narrative theorists.

---

6   Fernández (2018) offers an initial account of functionalism.

When Fernández's responds to these cases, however, his answers are effectively restatements of the mnemonic role. In cases of embellishment and isolation, the right dispositions are formed and maintained; In cases of confabulation and epistemic irrelevance, they are not. Nothing in his discussion engages with nor hinges on the details of the cases under consideration.

Consider the embellishment case, where *S*'s father shoots a white rabbit, and *S* later (mis)remembers the experience as one where his father shot a black rabbit. In addressing this case, Fernández's is not concerned with what happened during the events in question to lead to the misremembering. There is no discussion of how *S*'s perceptual and memory systems were working during the particular events of this case. We do not know, either during the original shooting or subsequent recollection, whether *S* was paying attention or tired or distracted, whether the event was significant or traumatic for *S*, or anything about what occurred to *S* in the interim between the shooting and the recollection. Fernández's is not concerned with the overall distribution of coloration in rabbits where *S* lives, nor with the father's hunting habits, nor *S*'s range of hunting experience or time with his father. Instead, the question of whether this counts as a case of remembering is pitched as a question about the general operations of *S*'s memory faculty —i.e., whether *S*'s perceptions typically cause mental images that result in episodic remembering.

This shift in perspective, from the particular case at issue to more general tendencies of memory, is intentional. It is meant to be a virtue of Fernandez's account. As he characterizes the key benefit of functionalism:

> On both versions [role and realizer] of functionalism, what matters for whether a subject is having a mental state of some type is not the causal relations that actually hold between that state and other mental states of the subject, as well as the subjects perceptual inputs and behavioral outputs, but the causal relations that tend to hold between all of those states (Fernández, 2019, p. 48).

The appeal to tendencies is meant to give us a perspective on the capacity as a whole, and to cut off the concern that a memory system must work perfectly in order to work properly. What matters for episodic remembering is not whether a person actually remembers the particular event they take themselves to be remembering now, but whether, in general, they remember events from their past. Appealing to the mnemonic role to characterize episodic remembering allows Fernández's to absorb cases where perceptual experiences mischaracterize the world and cases where the contents of memory degrade or go missing. Fernandez makes use of this leniency to absorb the embellishment cases like the rabbit hunt, as well as isolation cases.

The tension between this appeal to tendencies and the cases under consideration becomes clearer in the treatment of the cases that Fernández's wants to exclude from his account of episodic remembering —cases of confabulation and epistemic irrelevance.

As with the others, he introduces each puzzle with a detailed case. For confabulation, he describes a Korsakoff's patient who creates a false memory of riding on a train. For epistemic irrelevance, he describes a painter who fails to recognize a scene he is painting as being from a previous experience. His responses to these cases do not engage with the event details. Instead, the cases are excluded by stipulating that the person involved in each case has a deficit that renders them unable to produce mental states that play the requisite mnemonic role. The Korsakoff's patient has amnesia that interferes with memory's storage of perceptual images, which leaves him unable to form the needed kind of input tendencies. The painter has a different kind of deficit, which prevents stored mental images from being activated and endorsed as past experiences. This deficit renders the painter incapable of exhibiting the requisite form of output tendencies.

Effectively, Fernández's is claiming that a person who lacks the ability to retain mental images and/or put them to use in subsequent experience cannot remember because they lack a critical feature of the basic capacity. This point is fairly straightforward, and it is not particularly controversial. I suspect that both causal and narrative theorists would agree, but would not consider these cases relevant to their concerns in establishing causal or narrative conditions on remembering. Fernández's treatment of cases draws the distinction between remembering and its absence *across* individuals. A person either possesses a mental state type that plays this mnemonic role or they do not. Causal and narrative theorists, in contrast, look to draw the distinction *within* an individual, wanting to allow that a person could episodically remember in one case and not in another.

I illustrate this with evidence from causal theorists. Martin & Deutscher's (1966) version of the painter case, which serves as inspiration for Fernández's epistemic irrelevance, does not involve the stipulation of a cognitive deficit in the painter. We are left to suppose that, in general, the painter has a working episodic memory and tends not only to retain mental images from past experiences, but to deploy those images subsequently in ways that he recognizes as deriving from past experience. That the painter has a few such mental images which he fails to recognize as memories seems perfectly ordinary. I suspect all of us have at least a few such cases where this has occurred. It is these kinds of cases, which occur within the mental life of an otherwise capable rememberer, that Martin and Deutscher want to include in their account of memory. Viewing epistemic irrelevance in this way, it becomes difficult to see how Fernandez could exclude such cases. This painter's error does not derive from a general deficit; it's merely an image that misfires in a capacity for remembering that otherwise has the requisite tendencies.

A similar point can be made for the confabulation case. Fernández's focuses on a clinical case of confabulation, where the memory error occurs as part of a structural deficit, which in turn is symptomatic of a broader psychiatric disorder. The Korsakoff patient has amnesia that precludes the retention and reactivation of perceptual images. Clinical confabulation cases are interesting, but many philosophers of memory, including causal theorists especially (Bernecker, 2017; Robins, 2019) have been interested in

forms of confabulation that occur in 'everyday remembering' —i.e., in persons who do not have a clinical diagnosis, whose memories are functioning well or at least normally. Non-clinical confabulations are demonstrated in experimental contexts; most notably, in paradigms developed by Elizabeth Loftus and colleagues that use suggestive interviews to implant information that participants later interpret as remembered events (e.g., Loftus & Pickrell, 1995). Similar cases of confabulation can also be found in Martin and Deutscher's (1966) original discussion of the causal theory. Their view involves stipulating and then refining a causal condition on remembering. The causal constraint emerges out of consideration of the possibility of veridical confabulation —a case where one has an accurate representation of a past experience, but where the representation is brought about through some serendipitous string of events, not because of any connection to the past experience.

These are all cases where the person involved has a representation of a past event that has no connection to that past event. Importantly, though, there is no broader deficit suspected on the part of the person involved. Outside of the experimental paradigm, or without the convergence of a set of strangely serendipitous circumstances, the confabulation would not have occurred. More importantly, discussion of these cases proceeds against the backdrop assumption that the person's memory is otherwise working normally. In fact, Loftus uses these cases as a demonstration of the ways in which those of us with properly functioning memories can be susceptible to confabulation and false memory (Loftus, 2003).

Given that these confabulations occur *within* a properly or normally functioning memory, it seems that Fernandez would be compelled to include them. The overall tendencies that govern the mnemonic role are intact. This case of confabulation looks more akin to the case of embellishment. They're both cases where the memory system goes beyond what it should in a particular case, but otherwise is working well. Similarly, the case of epistemic irrelevance looks much like the case of isolation. Both are cases where the usual tendencies are lacking or weaker, but as anomalous instances are not worrisome. In summary: insofar as confabulation and epistemic irrelevance are understood as causal and narrative theorists interpret them, Fernández's view does not exclude them.

It's possible that Fernández's could develop this functionalist account further, in ways that could address these cases. Doing so, however, pushes against the nature of functionalism. The aim of the approach is to characterize the general role of a particular kind of mental state, and distinguish that role from the role of other mental states. Functionalism is not well suited to identifying or labeling a particular occurrent mental state as belonging to one type or another. It is not clear how functionalism could be used to classify a mental image. This was, in one sense, the criticism developed in the previous section.

To evaluate Fernández's functionalism as an account of episodic remembering, it would be good to see the view in its most fitting context —i.e., as an answer to the question Fernández's posed at the beginning of the chapter. How does episodic

Estud.filos n.º 64. Julio-diciembre de 2021 | pp. 201-222 | Universidad de Antioquia | ISSN 0121-3628 | ISSN-e 2256-358X 219

https://doi.org/10.17533/udea.ef.n64a11

remembering differ from perception, imagination, introspection, and the like? Fernández's does not answer this form of the question. The lack of an answer here is particularly worrisome for the view given the recent interest in *continuism* amongst philosophers of memory. Over the last two decades, philosophers and memory scientists have become increasingly interested in the relationship between episodic remembering and episodic imagining. There are now many who endorse the claim that remembering is a form of imagining (Hopkins, 2018; Michaelian, 2016) or that remembering and imagination are both forms of episodic simulation (Addis, 2018; 2020). These views are forms continuism. While there are others who defend discontinuism, arguing for memory and imagination to be kept distinct (Perrin, 2016; Robins, 2021), continuism is by far the more popular view (see Michaelian, Klein, & Szpunar, 2016). Without a discussion of how his view relates to these mental state proposals, it is difficult to determine whether the mnemonic role has been fully and fairly articulated.

Functionalism offers a popular and prominent account of mental states, and Fernández's attempt to use its tools to address the metaphysics of memory is innovative. Ultimately, however, it appears to be unsuccessful. Functionalism is not suited for answering the metaphysical question that has been of interest to philosophers of memory, so it is wrong to characterize it as a direct competitor to causal and narrative views. The view may still be useful in answering a different question. It may offer a unique way of accounting for the mnemonic role of episodic remembering. Given the extent of recent work on the connections between remembering and imagination, however, the uniqueness of this mnemonic role cannot simply be stipulated.

Conclusion

In this paper, I have presented a multi-dimensional critique of Fernández's functionalist account of the metaphysics of memory. I have argued that the mnemonic role, as he specifies it, leads to a number of challenges for the account and the judgments he wants to make about particular cases. I have also argued further that even if the mnemonic role is adjusted to address these problems, there are deeper problems with the use of functionalism to assess episodic remembering. Functionalism characterizes mental state types and is suited to assessments of episodic remembering *across* individuals, but the concerns of causal and narrative theorists involve token states of remembering and assessments *within* individuals.

Luckily for Fernández, who has structured his multifaceted account so that its various components are largely independent of one another, even if my assessment of this metaphysical proposal is correct, this does not threaten his account of the intentional and phenomenological aspects of episodic memory. Those aspects of the account, and the book, are exceptionally rich, and I look forward to future debates about how to craft a metaphysics of memory that suits them.

References

Addis, D. R. (2018). Are episodic memories special? On the sameness of remembered and imagined event simulation. *Journal of the Royal Society of New Zealand, 48,* 64-88. https://doi.org/10.1080/03036758.2018.1439071

Addis, D. R. (2020). Mental time travel? A neurocognitive model of event simulation. *Review of Philosophy and Psychology, 11*(2), 233-259. https://doi.org/10.1007/s13164-020-00470-0

Alvarez, M. (2017). Desire, dispositions, and the explanation of action. In J. A. Deonna & F. Lauria (Eds), *The Nature of Desire.* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199370962.003.0005

Bernecker, S. (2010). *Memory: a philosophical study.* Oxford University Press. https://doi.org/10.1093/obo/9780195396577-0072

Bernecker, S. (2017). A causal theory of mnemonic confabulation. Frontiers in theoretical and philosophical psychology. *Front Psychology, 8,* 1207. https://doi.org/10.3389/fpsyg.2017.01207

Brockmeier, J. (2015). *Beyond the archive: memory, Narrative, and the autobiographical process.* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199861569.001.0001

Debus, D. (2010). Accounting for epistemic relevance: a new problem for the causal theory of memory. *American Philosophical Quarterly, 47,* 17-29.

Fama, R., Patel, A. & Sullivan, E. V. (2012). Anterograde episodic memory in Korsakoff Syndrome. *Neuropsychology Review, 22,* 93-104. https://doi.org/10.1007/s11065-012-9207-0

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

Goldie, P. (2012). *The mess inside: narrative, emotion, and the mind.* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199230730.001.0001

Hopkins R. (2018). Imagining the past: on the nature of episodic memory. In F. Macpherson & F. Dorsch (Eds.), *Perceptual Imagination and Perceptual Memory* (pp. 46-71). Oxford University Press

Kopelman, M. D., Thomson, A. D., Guerrini, I. & Marshall, E. J. (2009). The Korsakoff syndrome: clinical aspects, psychology, and treatment. *Alcohol and Alcoholism, 44,* 148-154. https://doi.org/10.1093/alcalc/agn118

Leitgeb, H. (2017). *The stability of belief.* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198732631.001.0001

Loftus, E. F. (2003). Our changeable memories: legal and practical implications. *Nature Reviews: Neuroscience, 4,* 231–234. https://doi.org/10.1038/nrn1054

Loftus, E. F. & Pickrell, J. E. (1995). The formation of false memories. *Psychiatric Annals, 25,* 720-725. https://doi.org/10.3928/0048-5713-19951201-07

Martin, C. B. & Deutscher, M. (1966). Remembering. *Philosophical Review, 75,* 161-196. https://doi.org/10.2307/2183082

Mazzoni, G., Clark, A., De Bartolo, A., Guerrini, Ch., Nahouli, Z., Duzzi, D., De Marco, M., McGeown, W. & Venneri, A. (2019). Brain activation in highly superior autobiographical memory. *Cortex,* 1-36.

Michaelian, K. (2016) *Mental time travel: episodic memory and our knowledge of the personal past.* MIT Press. https://doi.org/10.7551/mitpress/10591.001.0001

Michaelian, K., Klein, S. B. & Szpunar, K. K. (Eds). (2016). *Seeing the future: theoretical perspecives on future-oriented mental time travel.* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780190241537.001.0001

Parker, E. S., Cahill, L. & McGaugh, J. L. (2006). A case of unusual autobiographical remembering. *Neurocase, 12,* 35-49. https://doi.org/10.1080/13554790500473680

Perrin, D. (2016). Asymmetries in subjective time. In K. Michaelian, S. B. Klein & K. K. Szpunar (Eds), *Seeing the future: theoretical perspecives on future-oriented mental time travel* (pp. 39-61). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780190241537.003.0003

Robins, S. K. (2020). Mnemonic confabulation. *Topoi, 39,* 121–132. https://doi.org/10.1007/s11245-018-9613-x

Robins, S. K. (2020). Defending discontinuism, naturally. *Review of Philosophy and Psychology, 11*(2), 469-486. https://doi.org/10.1007/s13164-020-00462-0

Schectman, M. (1994). The truth about memory. *Philosophical Psychology, 7,* 3-18. https://doi.org/10.1080/09515089408573107

Spiegel, D. R. & Lim, K. (2011). A case of probable Korsakoff's syndrome: a syndrome of frontal lobe and diencephalic structural pathogenesis and a comparison with medial temporal lobe dementias. *Innovative Clinical Neuroscience, 8,* 15-19.

Defending functionalism and self-reference in memory

Jordi Fernández

University of Adelaide, Adelaide, Australia
E-mail: jorge.fernandez@adelaide.edu.au

Abstract: In recent work, Sarah Robins, Gerardo Viera and Steven James have provided some insightful objections to the ideas offered in my book, *Memory: A Self-Referential Account*. In this paper, I put forward some responses to those objections. Robins challenges the idea that being a memory could be a matter of having a particular functional role within the subject's cognitive economy. Viera challenges the idea that the content of a memory could explain some of its phenomenological properties. And James challenges the idea that our memories could be immune to error through misidentification. All three commentators are targeting, not tangential aspects of, but fundamental assumptions in the account of memory proposed in the book. For that reason, modifying some of those assumptions would amount to proposing a whole different account of memory. I hope to show, however, that such a radical move is not necessary. For there are possible responses to the objections from all three commentators which are available within the constraints of the account proposed in the book.

Keywords: memory, functionalism, phenomenology, immunity to error through misidentification

Cómo citar este artículo

Estud.filos n.º 64. Julio-diciembre de 2021 | pp. 223-236 | Universidad de Antioquia | ISSN 0121-3628 | ISSN-e 2256-358X

223

https://doi.org/10.17533/udea.ef.n64a12

SIMPOSIO

Defendiendo el funcionalismo y la auto-referencia en la memoria

Resumen: En trabajos recientes, Sarah Robins, Gerardo Viera y Steven James han presentado algunas objeciones a las ideas ofrecidas en mi libro *Memory: A Self-Referential Account*. En este texto, presento algunas respuestas a esas objeciones. Robins desafía la definición funcional del recuerdo. Viera desafía la idea de que el contenido de una memoria pueda explicar algunas de sus propiedades fenoménicas. Y James desafía la idea de que nuestros recuerdos puedan ser inmunes al error de identificación errónea. Los tres comentaristas están apuntando, no a aspectos tangenciales, sino a supuestos fundamentales de la propuesta presentada en el libro. Por esa razón, modificar algunos de esos supuestos equivaldría a proponer una explicación de la memoria completamente diferente. Sin embargo, espero demostrar que un paso tan radical no es necesario: hay respuestas a las objeciones de los tres comentaristas que están disponibles dentro de las limitaciones de la teoría presentada en el libro.

Palabras clave: memoria, funcionalismo, fenomenología, inmunidad al error por identificación errónea

Jordi Fernández

is associate professor of philosophy at the University of Adelaide. He is the author of *Transparent Minds: A Study of Self-Knowledge* (2013) and *Memory: A Self-Referential Account* (2019).

ORCID: 0000-0002-4502-1003

First of all, I want to thank Sarah Robins, Gerardo Viera and Steven James for their insightful commentaries on *Memory: A Self-Referential Account*. I am not sure that I will be able to do full justice to all of their commentaries in my replies, but I hope that I have been able to concentrate on their most substantive points regarding the book. I am also very grateful to the editors of *Estudios de Filosofía* for giving us the opportunity to discuss the book *Memory* in this journal.

1. James on immunity to error through misidentification

Steven James (2021) concentrates on the discussion of immunity to error through misidentification (for short, 'EM') in the book, and my proposed defence of the view that memory judgments are IEM; a defence which appeals to the content of memories. James begins by motivating the view that memory judgments are not IEM through an intuitive example, and through an explanation of why we have the intuition that the memory judgment concerned is not IEM. Then, James puts forward one main criticism for my defence of the view that memory judgments are IEM. This is the criticism that the proposed defence trivialises the view. According to James (2021), the proposed defence of the view that memory judgments are IEM makes the view uninteresting, for two reasons. For one thing, it delivers IEM, as it were, on the cheap, because it is merely based on a technicality. For another, it deprives the thesis of any significance about self-consciousness.

Let us consider James's example first. In the example, Gayle recalls getting winked by Delta Burke. It turns out, however, that it was her sister Linda who got winked; an episode that Gayle witnessed, and one that she recalls correctly. In this example, James (2021) tells us Gayle's judgment that she got winked by Delta Burke is vulnerable to error through misidentification. For it is based on an accurate memory of a winking by Delta Burke. But Gayle wrongly assumes that the person who got winked was her and, as a result, her judgment that she got winked by Delta Burke is wrong. Does this not show that memory judgments are not IEM? It seems to me that this example trades on a certain ambiguity about the kind of mental image involved in Gayle's remembering. Gayle may be visualising the winking episode either from the first-person point of view (or from the inside), or from the third-person point of view (or from the outside, with her as one of the participants in the episode). In the former case, I would argue that, even if Gayle's mental image qualifies as a memory, it is not an accurate memory. After all, the remembered episode did not happen as Gayle is visualising it (Delta did not wink while looking at Gayle). In the latter case, I would argue that Gayle's mental image does not qualify as a memory to begin with.

According to James (2021), if we assume the functionalist account of the metaphysics of remembering proposed in the book, then Gayle's mental image of the winking episode should qualify as a memory. For it is the kind of thing that tends to be caused by such

an event happening. This may be true if Gayle is visualising the winking episode from the first-person point of view, but not if she is visualising it from the third-person point of view. In the latter case, it is not the kind of mental image which tends to be caused by winking episodes. The reason is that Gayle is never in a position to perceive herself being winked at, at least not in the manner in which a witness of the episode would be in a position to perceive Gayle when somebody winks at her. Arguably, she can never be in such a position, because she cannot be in two places at once.

The second horn of my response reveals a deeper disagreement between James and me. According to James (2021), it is no wonder that, in the Delta Burke example, we have the intuition that Gayle's judgment is vulnerable to error through misidentification. After all, we know that memories are formed, modified over time, and reconstructed later (often repeatedly). The way James sees it, then, it would be surprising if memory processes did give rise to the IEM phenomenon. And I can see how if someone accepts such a liberal conception of what counts as a memory, then they would reach the conclusion that memory judgments are not IEM. I therefore think that, ultimately, the disagreement between James and me, with regards to the IEM phenomenon, stems from a difference in our conceptions of what it is for someone to remember something. On the functionalist construal of remembering, a mental image of some event does not need to have been caused by a perception of that event in order for it to count as a memory of the event (if it did, we would not be able to misremember events that we correctly perceived in the past). Nevertheless, on the functionalist construal of remembering, the mental image does need to be the type of mental image which would normally be caused by such an event. And the reconstruction processes to which James (2021) alludes can interfere with this property of the mental image. In the Delta Burke case, for instance, Gayle has a mental image which has been reconstructed from her original perception of the winking episode (assuming, that is, that Gayle is visualising the winking episode from the third person point of view). And the degree of reconstruction is such that the resulting mental image is no longer the kind of mental image which would normally be caused by winking episodes, because we never perceive ourselves from the outside. In this case, Gayle's memory of the event has not been reconstructed by the processes to which James is referring. The way I see it, Gayle's memory has been eliminated by those processes.

Let us now consider the criticism that my proposed defence of the view that memory judgments are IEM makes the view uninteresting because it delivers IEM on the cheap. The reason why James thinks this is that my definition of what it is for a memory judgment to count as being IEM requires that it is impossible for a misidentification error to occur while the memory on the basis of which the judgment has been made remains fully accurate. This condition is doing the heavy lifting in my defence of the IEM view. However, James (2021) objects, if we include this condition in our definition of what it is for a memory judgment to be IEM, then IEM becomes not only a matter of whether it is impossible for the judgment to be false because of a misidentification

of the person, or object, about whom one made the judgment. And, the implication is, IEM should only be a matter of whether that possibility has been ruled out. Thus, the reason why my proposed defence of the IEM view succeeds, James suggests, is that we have re-defined the IEM notion in a somehow inappropriate, or ad hoc, way.

I disagree with the claim that if we require, for a memory judgment to be IEM, that it is impossible for a misidentification error to occur while the memory on the basis of which the judgment has been made remains fully accurate, then this definition makes the thesis that memory judgments are IEM uninteresting. On the contrary, it seems to me that the thesis becomes uninteresting if we do not include such a requirement. For if we do not include such a requirement, I agree with James (2021), the thesis becomes false. But it becomes trivially false. Suppose that I misremember being in front of a tree a week ago and, on the basis of that incorrect memory, I judge that I was in front of a tree. If the question of whether my judgment is IEM is only the question of whether it is possible for my judgment to be wrong because I was not in front of a tree, even though someone else was, then, naturally, my judgment is not IEM. Someone else could have been in front of the tree that I am misremembering. But, then, the reason why my memory judgment is not IEM is simply that memory is fallible. At any point, any of my memories may falsely present to me an event consisting in my instantiation of some property in the past. If, in the past, that property happened to be instantiated by some other person, then the possibility that IEM is supposed to rule out will obtain. This reading of the claim that memory judgments are not IEM, however, only tells us that memory is fallible, and surely that is not the interest of the IEM thesis with regards to memory. The reason why we need the requirement that, in order for a memory judgment to be IEM, it needs to be impossible for a misidentification error to occur *while the memory on the basis of which the judgment has been made remains accurate* is precisely that we want to rule out the possibility that an error occurs just because memory is fallible.

Let us consider, finally, the criticism that my proposed defence of the view that memory judgments are IEM makes the view uninteresting because it deprives the thesis of any significance with regards to self-consciousness. In James's view (2021), the proposed defence of the IEM view tells us that the reason why memory judgments are IEM is that memory judgments have a certain content, a content which involves the self. And, from this diagnosis, I draw a lesson with regards to self-consciousness, namely, that our first-person conception of ourselves, or self-conception (a conception formed through faculties such as introspection, proprioception, or memory), includes the fact that we are the kinds of beings which are not only extended in space, but also in time. Now, if this is right, James objects, then the view that memory judgments are IEM does not tell us anything interesting about our self-conception. The view that memories have certain contents which involve the self does tell us something interesting about our self-conception, but the view that memory judgments are IEM plays no role in delivering the relevant result.

I confess that I am a little confused about the dialectic here. James (2021) seems to concede that the fact that memory judgments are IEM suggests that memories have certain contents which involve the self. And he also seems to concede that the fact that memories have certain contents which involve the self suggests that our self-conception is that of a being extended in time. Then, why does he dispute the claim that the fact that memory judgments are IEM suggests that our self-conception is that of a being extended in time? If we assume that 'suggests that' is transitive, James's worry cannot really be that the view that memory judgments are IEM plays no role in delivering the result that our self-conception is that of a being extended in time. Instead, his worry must be that this is a result at which we do not need to arrive through the IEM route. If we have other grounds for believing that memories have certain contents which involve the self, then we will obtain the same result about our self-conception, independently of our position with regards to the question of whether memory judgments are IEM. If this is James's point, then I do not disagree with his point. But it seems to me that there is a considerable leap, from this point, to the point that the IEM phenomenon in memory tells us nothing interesting about our self-conception.

2. Viera on the feeling of pastness

Gerardo Viera (2011) concentrates on the discussion of the feeling of pastness in memory, and my proposed explanation in terms of it; an explanation which appeals to the content of memories. Viera raises two concerns for my explanation, depending on how exactly the explanatory link between the content of memories and the feeling of pastness is conceived. He takes these two concerns to show that the proposed explanation does not succeed. Accordingly, he proposes a different way in which time can be part of the content of memories, in terms of what he calls 'path-dependent representations' of time. Viera's proposal makes use of David Lewis's distinction between personal time and external time. I think that this distinction is very helpful for the purposes of explaining the feeling of pastness in memory, and I am grateful to Viera for bringing it up. As far as I can see, however, this distinction can actually be deployed to address Viera's two concerns about the explanation of the feeling of pastness in memory in terms of memories having self-referential contents.

If I understand Viera's (2021) discussion of my proposed explanation of the feeling of pastness correctly, his criticism of it has the form of a dilemma. My proposal is that, when we are aware of a remembered event as being in the past, what the relevant memory represents is that a perception of that event has caused the memory. What the memory represents, then, is a certain property of the remembered event, namely, the property of being at the causal origin of the memory. The question that Viera seems to be raising, in response, is the question of whether that property is the property of being in the past, or not. And, the way he sees it, significant difficulties are going to

arise for my proposed explanation of the feeling of pastness whether we answer this question in the affirmative, or in the negative.

If we decide that the property of being at the causal origin of a memory is not the property of being in the past, then the proposed explanation is an error theory. The outcome is that the feeling of pastness is not a feeling of pastness, but a feeling of causation. We misdescribe this experience as a feeling of pastness but, due to the close connection between causation and time, our experience of a remembered event always picks up an event which, in fact, is in the past.[1] This, Viera (2021) tells us, is not a satisfactory explanation of the feeling of pastness, and it is not what my proposed explanation of the feeling of pastness promised us. By contrast, if we decide that the property of being at the causal origin of a memory just is the property of being in the past, then my proposed explanation of the feeling of pastness becomes similar to an objectivist approach to colour. The thought is that, by looking at a blue wall, my perceptual experience represents what is in fact a reflectance property of the surface that I am looking at, and colour blue is simply that property. Analogously, by remembering my seeing an apple, my memory experience represents a property of that event, namely, the property of being at the causal origin of my memory. And being in the past is simply that property.[2] But this, Viera tells us, is highly implausible. The pastness of a remembered event cannot consist in a causal property of it.

In fact, Viera could have pointed out that things seem even worse for the proposed account of the feeling of pastness in terms of memories having self-referential contents. It is not only that the account seems to run into trouble whether the crucial question of whether, in memory, we represent the past is answered in the affirmative or in the negative. It is also that, in different places, I seem to be giving different answers to the same question. So the account, one could argue, is not only uninformative, but also incoherent.

In order to present his own account of the feeling of pastness, Viera raises Lewis's distinction between external time and personal time. Normally, a remembered event is earlier than the relevant memory in both personal time and external time. In Viera's example, he remembers, in 2021, his struggling to swim in 1991. His struggle, then, is earlier than his memory of it in external time (since 1991 is earlier than 2021), and in personal time (since it is part of Viera's history, or his life). The case of the time traveller, however, illustrates how a remembered event can be earlier than the relevant memory in personal time, but not in external time. If Viera travels back to 1989 and he remembers, then, his struggling to swim in 1991, then his struggle is earlier than his memory in personal time (since the event remains a part of Viera's life) but not in external time (since 1991 is not earlier than 1989). Viera's suggestion, then, is that, in

---

1   See (Fernández, 2019, pp. 108-109) for this view.

2   See (Fernández, 2020, p. 294) for this view.

memory, we are aware of an event as being in the past because our memory of it gives us a variety of cues to locate the event in our personal past.

Viera's account of the feeling of pastness is an interesting proposal. However, I will not be engaging with it as a competing account. Instead, I would like to show that, once we help ourselves to the distinction between personal and external time, we can tackle Viera's dilemma for the account of the feeling of pastness in terms of memories having self-referential contents. The question, let us remind ourselves, is whether the property of a remembered event of being at the causal origin of the relevant memory is the property of being in the past or not. I am inclined to give different answers to this question depending on whether we have external time or personal time in mind.

If we have external time in mind, then the answer is 'no'. In that case, I agree that my account of the feeling of pastness becomes an error theory. We talk of a feeling of pastness but, strictly speaking, what the feeling is a feeling of is a causal relationship holding between the memory that we are having and the remembered event. Viera replies that this is not how the proposed account of the feeling of pastness is presented, but I think I have been quite explicit about what the proposed view, both in the book, and in earlier work.[3] Leaving aside issues of presentation, though, Viera's (2021) concern may be that an error theory of the feeling of pastness is not much of an explanation. In that case, I have little to say in response. I can point out that this error theory tells us, not only why it is not appropriate to talk of a 'feeling of pastness' in memory, but it also explains why we are inclined to talk in that way. But if the objection concerns the deflationary nature of error theories as explanations generally, then I concede to Viera that the proposed account is an 'explanation' of the feeling of pastness in a weak sense of the term.

If we have personal time in mind, then I am inclined to answer 'yes' to the question of whether the property of a remembered event of being at the causal origin of the relevant memory is the property of being in the past. In this case, I embrace the analogy with colour objectivism. The feeling of pastness is the way in which we experience something about remembered events that our memories represent. What they represent is not that those events have some position in external time, but the fact that those events are causally related to our memories. And the property of being causally related to our current states, such as our memories, is the property of being in our personal past. Think about it in this way: What is it for a remembered event to be in our personal past, that is, to be part of our life? It is for it to be something that happened to us. And what is it for an event to be something that happened to us? It is for the person to whom it happened to be identical with the person who is having the relevant memory. But what makes the two people identical? One plausible answer, it seems to me, is that the properties of one of the two people are causally responsible for the properties of the

---

3   See footnote 2, and (Fernández, 2008, p. 349).

other one. And, in that case, it seems reasonable to think that what it is for an event to be in our (personal) past is for it to be causally related to our current properties. If the event in question is a remembered event, then our memory of it seems the natural candidate for our relevant current property.

In response to the colour objectivism approach, Viera objects that this approach commits us to the prediction that, in other instances in which we represent causation by occupying other mental states, then we should find that we experience a similar phenomenology to the phenomenology of pastness in memory. And this prediction, Viera points out, seems to be wrong. It seems to me, however, that the approach only commits us to the prediction that, in other instances in which we represent causation in the way represented by memories by occupying other mental states, then we should find that we experience a similar phenomenology to the phenomenology of pastness in memory. And this prediction seems to be consistent with the empirical evidence provided by Viera. For, as far as I can see, the mental states involved in the cases presented by Viera, are not causally self-referential. If they represent causation, they do not represent it by representing themselves as one of the relata involved in the represented causal relations.[4]

Viera's own account of the feeling of pastness in terms of path-dependent representations of time may be a compelling alternative to the subject-dependent and subject-independent theories of temporal representation that I consider, and ultimately dismiss, in the book. My aim in this discussion has not been to challenge Viera's account, but to suggest that, even if it is successful, my proposed account of the feeling of pastness, in terms of memories having self-referential contents, can overcome the difficulties that Viera raises for it, assuming the same conceptual resources.

3. Robins on Functionalism

Sarah Robins (2021) discusses the functionalist account of remembering proposed in the book, and she raises four worries for the account. Two of those worries are general worries, and two of them are more specific ones. The first general worry is the following: on the functionalist account, having a mental image i counts as remembering that some proposition p is the case if and only if having i has a certain functional role in the subject with regards to the fact that p. But having i, Robins points out, is an occurrent mental state and, as such, it can have no functional role. It does not tend to be produced by anything, and it does not tend to produce anything. The two specific concerns about the functionalist account of remembering stem from this general worry. Robins (2021)

---

4   I therefore disagree with the view (Searle, 1983) according to which perceptions are causally self-referential. Our intuitions about veridical hallucination seem to suggest that the kinds of thought experiments employed to motivate the causally self-referential nature of memories would not work in the case of perception.

claims that the functionalist treatment of the Embellishment case, and the functionalist treatment of the Confabulation case, are not satisfactory due to this shortcoming in the functionalist account. Robins considers various readings of the claim that having mental image i plays a functional role in the subject with regards to the fact that p, and concludes that none of them can deliver a satisfactory account of our intuitions about remembering in both Embellishment and Confabulation. The second general worry is that, since having mental image i is an occurrent mental state, no reading of the claim that having i has a certain functional role with regards to the fact that p can capture our intuitions about what it is to remember that p. For functional roles concern causal relations in which types of mental states, such as mental images, are involved. And what we want to know, when we ask what it is for a subject to remember that p in virtue of having mental image i, is what it is for the subject, in this particular instance of their having i, to remember that p. Our question, in other words, concerns the token mental state that the subject occupies, and not the type to which it belongs.

Let us take the first general concern first. Robins's concern is well-taken in that, sometimes, I speak of tendencies in which having a mental image i is involved, and sometimes I speak of the episode of having mental image i satisfying certain counterfactuals. I do waver between these two conceptions of what it is for mental image i to have a certain functional role, which Robins has noticed. To simplify matters, then, let us focus on the counterfactual reading of the functionalist account, and see whether that counterfactual reading can give us a successful treatment of both Embellishment and Confabulation. The idea is that a subject S having a mental image i remembers that p just in case, if S perceived that p, then they would have a mental image of i's type. And if they had a mental image of i's type, then they would believe both that p and that they perceived that p in the past. There is an appeal to types here.[5] Which is, then, the relevant type? I am inclined to individuate mental images phenomenologically. Thus, we can think of the relevant type as the phenomenological type to which mental image i belongs.

In Embellishment, the intuition is that the subject remembers a black rabbit being shot (even though, in fact, a white rabbit was shot). And it seems to me that the relevant counterfactual is indeed satisfied: If the subject had witnessed a black rabbit being shot, the mental image that they are actually having is the type of mental image that they would have had. To be sure, they remember it incorrectly because, even though the mental image they are having is the type of mental image that they would have had if they had witnessed a black rabbit being shot, they did not in fact witness a black rabbit

---

5   Robins is right in pointing out that the functionalist account of remembering makes a further appeal to a background of normal circumstances. Thus, the counterfactuals above would need to be qualified by prefacing them with 'normally'. However, I do not believe that this qualification is necessary in order to deal with the cases of Embellishment and Confabulation, which are the cases on which Robins is concentrating. Thus, for the purposes of this discussion, I will leave the qualification implicit in the formulation of those counterfactuals.

being shot. Nevertheless, the counterfactual is satisfied, which captures, I suggest, the intuition that the subject is (incorrectly) remembering that a black rabbit was shot. Robins (2021) objects that the fact that, in Embellishment, the counterfactual is satisfied does not allow us to reach any conclusions about the tendencies that the mental image of the black rabbit being shot has in the subject. This seems right. But I am not offering, here, the counterfactual reading of the functionalist criterion for remembering as a way of specifying what the tendencies of mental images are in remembering. I am offering it as a way of specifying what their functional roles amount to.

In Confabulation, the intuition is that the Korsakoff subject does not remember having a conversation in a train. And it seems to me that the counterfactual is not satisfied: If the subject had experienced the conversation in the train, the mental image that they are actually having is not the type of mental image that they would have had. The reason, I claim, is that the Korsakoff subject has amnesia. Now, Robins objects that, even though Korsakoff patients typically have amnesia for events which happened after the development of their disorder, they do not need to have amnesia for events which happened before the development of their disorder. That is, they do not need to lose memories formed prior to that point. Thus, Robins suggests, we can imagine a case of confabulation which involves the report of a conversation in a train happening before the development of the subject's disorder. And the functionalist account, Robins tells us, will not be able to rule out such a confabulation case as a case of remembering since, in that case, the mental image that the subject is having will satisfy the required functional role. With regards to this objection, my response is, in a way, conciliatory and, in a way, adversarial.

Let us take the adversarial part of the response first. I would like to cast doubt on the idea that, in the confabulation case imagined by Robins, the subject's image does satisfy the required functional role. Is it the case that, had the subject been in a train conversing with someone, they would now have the type of mental image that they are actually having? If the possible conversation happens before the development of the subject's disorder, the answer is 'yes'. But if the possible conversation happens after it, then the answer is 'no'. It seems, then, that the mental image that the subject is having plays the functional role required for the subject to remember that they had a conversation in a train *at such-and-such time* (a time which is, in fact, previous to the development of their disorder, whether they can identify it as such or not), but it may not play the functional role required for the subject to remember, more simply, that they had a conversation in a train.

But isn't this all Robins needs in order to make her point? After all, Robins's point was that a patient with Korsakoff's syndrome, and a particular form of amnesia, could still count as remembering, which seemed to conflict with the functionalist approach to memory. And I am granting that, in the case envisaged by Robins, the patient has a mental image which has the required functional role for the fact that they had a conversation in a train at such-and-such time (a time which is, in fact, previous to the

development of their disorder). One might think, therefore, that this case does present a challenge to the functionalist framework, since we are granting that, depending on the particular details of the patient's amnesia, the case could indeed count as a case of remembering.

The conciliatory part of the response is to concede that not all cases of confabulation *need* to be ruled out as cases of remembering by the advocate of the functionalist approach. If there are cases in which the subject is confabulating when they have their mental image, but their mental image does satisfy the necessary functional role, then I have trouble seeing why we should rule them out as cases of remembering. To be clear, my reason for ruling out Confabulation as a case of remembering was not that the case involves confabulation ('Confabulation' then, may have been a misnomer), but that the case involves the type of amnesia which interferes with the mental image satisfying the functional role which is required for remembering. If there is a type of amnesia which is consistent both with confabulation and with the mental image that the subject has satisfying the necessary functional role for remembering, then I am willing to accept that some cases of confabulation are cases of remembering. At the very least, we should ask what independent grounds we have for assuming that memory and confabulation need to be incompatible.

Robins's (2021) second general worry is that, since having mental image i is an occurrent mental state, the appeal to functional roles cannot capture our intuitions about what it is to remember that p in virtue of having i. For functional roles concern causal relations in which types of mental states are involved, and what we want to know, when we ask what it is for a subject to remember that p in virtue of having mental image i, is what it is for the subject, in this particular instance of their having i, to remember that p. Now, it seems to me that if a specific mental state, which is occupied by a specific subject at a specific time, is a particular instance of a mental type, such as remembering that p, then the subject qualifies as remembering that p at the relevant time. For that reason, it seems to me that if we want to know, for a specific, occurrent mental state, whether it counts as remembering that p, it is a legitimate question to ask whether the mental state is of a certain type, namely, remembering that p. The type of mental state will provide us with the conditions that the occurrent mental state needs to satisfy. I do not see any confusion here.

According to Robins, the causal theorist of memory, for example, and myself are talking past each other, since the causal theorist is concerned with token mental states whereas I am concerned with types of mental states. I disagree. Both of us are concerned with token mental states. I am just answering the question of whether the token mental state in question is a state of remembering by considering whether it belongs to a certain type, a type which requires certain conditions. Whether the causal theorist is doing the same thing or not will depend on their approach to causation. If they regard the kind of causation involved in memory as token causation, then they will answer the question in a different way. But if they do not regard it as an instance of token

causation, they will also consider whether the token mental state under consideration belongs to a certain type. It is just that the conditions required to belong to that type will be different from mine (they may involve probabilities, processes, or other factors, but they will be, in any case, backward-looking conditions).
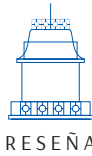
Let us keep in mind that the functionalist account of remembering is not meant to be an account of remembering correctly. It is meant to be an account of remembering simpliciter. Thus, the thought is that, when we ask what it is for a subject to remember that p in virtue of having mental image i, we are asking a question about the particular mental image i, but the answer depends on facts about the type of mental state to which mental image i belongs. By contrast, when we ask whether the subject is remembering p correctly in virtue of having mental image i, we are also asking a question about the particular mental image i, and facts about the type of mental state to which mental image i belongs will not provide us with an answer to that question. If what the causal theorist and the narrative theorist are after is an account of remembering correctly, then the functionalist approach will not help them to build such an account. I agree with Robins on that point.

4. Concluding remarks

There is an interesting common factor in some of the objections raised by James (2021), Viera (2021) and Robins (2021). The common factor is that all three commentators are targeting, not tangential aspects of, but fundamental assumptions in the self-referential account of memory proposed in the book. Some of James's objections, for example, hinge on whether memories can be reconstructed or not. Likewise, Viera's objections depend on the issue of how intentionalist explanations of phenomenology work, and whether the components of the content of a mental state can account for what it is like for one to be in that mental state. These issues are not discussed in the book. They are mentioned, in the first chapter, but they are only mentioned as assumptions. There is a sense, then, in which all three commentators are focusing on features which are essential to the self-referential account of memory offered in the book. I do not believe that the account can therefore be modified in order to address the relevant concerns. Instead, the account becomes a different account of memory if some of those non-negotiable assumptions are dropped. However, I hope that I have made a cogent enough case for the idea that, on reflection, dropping the relevant assumptions will ultimately not be necessary.

References

Fernández, J. (2007). Memory and time. *Philosophical Studies, 141,* 333-356. https://doi.org/10.1007/s11098-007-9177-x

Fernández, J. (2019). *Memory: A Self-Referential Account.* Oxford University Press. https://doi.org/10.1093/oso/9780190073008.001.0001

James, S. (2021). Immunity to error through misidentification and the functionalist, self-reflexive account of episodic memory. *Estudios de Filosofía, 64,* 189-200. https://doi.org/10.17533/udea.ef.n64a10

Robins, S. (2021). The failures of functionalism (for memory). *Estudios de Filosofía, 64,* 201-222. https://doi.org/10.17533/udea.ef.n64a11

Searle, J. (1983). *Intentionality.* Cambridge University Press. https://doi.org/10.1017/CBO9781139173452

Viera, G. (2021). Feeling the past: beyond causal content. *Estudios de Filosofía, 64,* 173-188. https://doi.org/10.17533/udea.ef.n64a09

R E S E Ñ A

Fernández, J. (2019). *Memory: a self-referential account.* Oxford University Press

Juan F. Álvarez

Université Grenoble-Alpes, Grenoble, France
Email: jalvarces@gmail.com | ORCID: 0000-0002-3726-470X

Fernández' most recent book constitutes an articulated development of several philosophical considerations on memory displayed in his previous and forthcoming publications. The result of such an articulated development ends up being a consistent account that provides an innovative and thought-provoking perspective on episodic remembering. This volume not only gathers and articulates the author's previous ideas, but also provides new reflections on, and objections to alternate theories of memory, which encompass four significant domains in the philosophy of memory. In the first part of the book (Chapters 1, 2, and 3), Fernández offers an account of both the metaphysics and the intentionality of episodic memory; in the second part (Chapters 4 and 5), the author deals with certain phenomenological aspects involved in remembering; in the third part (Chapters 6 and 7), two important debates in the epistemology of memory are discussed.

Chapter 1 spells out the *explanandum* of the book, and the adopted strategy to construct a suitable *explanans.* The metaphysical, intentional, phenomenological, and epistemic aspects of episodic memory compose the project's *explanandum.* The first aspect is related to the conditions under which a mental state qualifies as an instance of episodic remembering. The second one concerns the representations or contents of memories. The third aspect is about the experience or what-it-is-like aspect of remembering. The fourth one is related to the kind of justification that episodic memory provides for knowledge. Fernández' strategy consists in taking both

Estud.filos n.º 64. Julio-diciembre de 2021 | pp. 237-243 | Universidad de Antioquia | ISSN 0121-3628 | ISSN-e 2256-358X

https://doi.org/10.17533/udea.ef.n64a13

237

the intentionality and the metaphysics of memory as fundamental pillars on the basis of which the explanations about the phenomenological and the epistemological issues are constructed. More specifically, the intentional pillar —what the author calls the "self-referential view of mnemonic content"— is the principal source of information used in the accounts of both the experience and the epistemic justification afforded by episodic memory. Fernández' self-referential view is actually the main pillar of the book because his accounts of the phenomenology and the epistemology of memory depend on it, whereas the metaphysical pillar simply works as a background that provides the definition of what a memory is. Both pillars are then independent: if Fernández' metaphysical view turns out to be false, the self-referential view may remain true; if the latter turns out to be false, the former may remain true.[1]

The construction of the metaphysical pillar is the goal of Chapter 2. Fernández uses a functionalist framework, according to which episodic memories are determined by their causal relations with certain inputs and outputs of the subject's cognitive system. Instances of remembering are thus individuated not in terms of some intrinsic property, but in terms of the functional role the state plays. A mental state qualifies as an instance of remembering, i.e., plays memory's functional role only if: first, the mental state tends to be caused by a specific input, which is the past experience of the event; second, the mental state tends to cause a certain output in the subject, namely, a disposition to believe both that the remembered event occurred and that she experienced it. Accordingly, if a mental state of the subject's cognitive system meets these two conditions, then the mental state in question is playing a mnemonic role; and if the subject's cognitive system has a mental state that plays a mnemonic role, the subject is remembering.

The intentional pillar is built in Chapter 3. Fernández' proposal on the intentionality of memory is the result of a rigorous intuition-based analysis of the truth conditions of mnemonic content. According to the self-referential view defended by the author —which he calls the "reflexive view"— memories represent themselves as coming from a veridical perception of the fact that the subject affirms to remember, i.e., memories represent their own causal origin. Given the reflexive view of the self-referential approach, at least four elements are involved in mnemonic content: (i) the memory itself, (ii) the veridical past perceptual experience, (iii) the objective fact, and more importantly, (iv) the relations involving those elements, some of which are causal relations. Moreover, the content of a memory could be spelled out in the form of a proposition that connects these elements in a suitable way, such as: 'the memory in question is caused by a subject having perceived an objective fact through a perceptual experience of it' (*Cf*. Fernández, 2006, p. 54; 2019, p. 79).

1　However, if this is the case —if the self-referential view turns out to be false— Fernández' proposals on the phenomenology and the epistemology of memory will also be false.

Since Fernández assumes that the intentional properties of memories are responsible for their phenomenal properties, i.e., the phenomenal properties depend on the intentional properties, his self-referential view constitutes the account's core of the phenomenology of remembering. Fernández' account focuses on three phenomenological features that compose the what-it-is-like aspect of memory: two of those features are related to the experience of time, and the other one to the feeling of ownership. Chapter 4 provides the explanation of the former features, which are the "awareness of previous experience" —the awareness of what it was like for the subject to perceptually experience the remembered fact— and the "feeling of pastness" —the awareness of the remembered fact as having obtained in the past. From the self-referential point of view, both experiences are due to the kind of content carried by memories. On the one hand, the presence of the awareness of previous experience is due to the experience of component (ii) of the mnemonic content —the veridical past perceptual experience. In other words, in virtue of the subject's memory representing the relevant past perceptual experience, the subject is able to have the awareness of the qualitative properties involved in the past perceptual experience. On the other hand, the presence of the feeling of pastness is due to the experience of component (iv), which in this case is the causal relation between the past perception of the fact and the resulting memory. Thus, by representing and experiencing this past causal relation the subject is able to subsequently experience her memory as having obtained in the past. An important consequence of this idea is that the feeling of pastness is not the awareness of a fact's temporal property, but the awareness of the memory's causal origin, which is, again, part of its content.

Chapter 5 offers the explanation of the phenomenal feature related to the sense of "mineness" that memory seems to involve, which is characterized as the "feeling of ownership" or the subject's experience that she is the owner of her memories. The nature of this experience is allegedly explained by what the author calls the "endorsement model", according to which a subject endorses her memories because they seem to match the past. Subjects are then aware of their memories as being their own to the extent that they are aware of them matching the past. Once again, Fernández' endorsement model is based on the self-referential view, given that the model takes a component of mnemonic content in order to explain the presence and absence of the feeling of ownership. In this case, the relevant component is (ii) —the veridical past perceptual experience. In particular, the issue here is related to the perceptual experience's veridicality. In normal circumstances, a subject can endorse a memory and claim that it matches the past because it is assumed that the experience was veridical. If the subject brings into question the veridicality of a remembered perceptual experience, she cannot have the distinctive phenomenology of her memory as being her own.

In Chapter 6, Fernández' thesis is that, due to the nature of mnemonic content, memory judgments have a particular epistemic aspect: immunity to error through misidentification, whereby one cannot be wrong as to the identity of the remembered subject of the experience. A subject's memory judgment may constitute an error through misidentification when three conditions are fulfilled: (a) the memory represents a subject as having had certain property, (b) the memory is fully accurate, and (c) the subject mistakenly thinks that she is identical to the represented subject. Although the very existence of observer memories —memories that show the subject as part of a remembered scene— might undermine Fernández' thesis because the subject could easily misidentify herself, the author argues that component (ii) of mnemonic content —the past perceptual experience— is responsible for preventing memory judgements fulfilling condition (c).[2] More precisely, a further analysis of component (ii) allegedly shows that part of the content of the remembered perceptual experience is the self as the bearer of extrinsic properties, i.e., in a remembered perception subjects are aware of themselves as the bearers of certain relations to objects of the perceived scene. Thus, if the remembered perceptual experience implies self-awareness of the experiencer, she cannot be mistaken in thinking that she is identical to the represented subject.

Finally, Chapter 7 offers a new position in the debate on the epistemology of memory between preservativism and generativism. The central question of this debate is whether memory merely *preserves* epistemic justification or can also *generate* it. Based on the self-referential view, Fernández makes a case for a form of generativism according to which memory is a basic epistemic source, that is, memory is an independent source that generates justification without relying on other epistemic sources. What allows the author to argue for this view is component (iv) of mnemonic content —the causal relation between the memory itself originating from a veridical perception of a fact. Since (iv) is only provided by memory and not by any other source, subjects can have justification for forming beliefs about the causal histories of their own memories. Note that not even the relevant perception in which the memory state originated can afford this justification because, at the moment of perceiving, the content that the eventual memory causally originated in the current perception is not available. Only episodic memory generates this content, which can constitute new grounds to form certain justified beliefs.

---

2   Fernández' discussion of observer memories is, however, more intricate. Although for the author it is fairly clear that the possibility of having observer memories does not undermine the fact that memory is immune to error through misidentification, the specific reason that supports this claim may vary. After examining three variants of the same observer memory case, the author provides three possible reasons that might support such an idea: observer memories do not engender false memory judgements, or they are not fully accurate, or they are not genuine memories (Fernández, 2019, p. 156). Fernández nevertheless is not explicitly committed to one of these reasons, and for him the possibility that one of them might potentially explain why having observer memories does not undermine the fact that episodic memory is immune to error through misidentification seems to be enough.

There are at least two main contributions to the philosophy of memory that Fernández makes that are worth highlighting. Firstly, Fernández' functionalist theory enriches the discussions on the metaphysics of memory that are dominated by causal and simulation theories nowadays. As noted by Michaelian & Robins (2018), given that memories —under the functionalist theory— do not need to be actually caused by past experiences, Fernández' metaphysical proposal denies both the sufficiency and necessity of a causal connection between current memories and past experiences in order to account for remembering. Therefore, along with the simulation theory, the functionalist explanation represents a rupture with the predominant causal account in the philosophy of memory (Michaelian & Robins, 2018, p. 28).[3] Secondly, Fernández' proposal on mnemonic content as a source of theoretical tools to clarify other philosophical issues may have interesting explanatory potential. On the one hand, philosophers could use this argumentative strategy to propose alternative perspectives to other recent debates in the philosophy of memory. On the other hand, philosophers could criticize and reject the strategy, its explanatory value, and its implications for the phenomenology and the epistemology of memory. Either option should provide interesting and important philosophical reflections on the nature of episodic memory.

Philosophers of memory already started to take into consideration Fernández' proposals with a critical spirit. Sant'Anna & Michaelian (2019) and Bernecker (2020) have shown some difficulties in Fernández' accounts of the metaphysics and the intentionality of remembering. Sant'Anna & Michaelian allege that the functionalist theory is too strict because it rules out both cases in which the subject does not form the disposition to believe that the remembered event occurred, and cases in which she forms instead the disposition to disbelief the occurrence of the event. They also claim that, since the theory rejects the necessity of a causal connection between memories and past experiences, it has difficulties in explaining the alleged particularity of remembering, which is that memories are about particular past events.[4] In addition, Bernecker has pointed out, in the first review of the book, that the self-referential view of mnemonic content is mysterious and needs further explanations, because it is unclear how different memories represent their distinct causes when this information is inaccessible to the rememberer.

Regarding the phenomenology and the epistemology of memory, Perrin, Michaelian & Sant'Anna (2020) and Michaelian (2020) criticize some aspects of Fernández' proposal. Perrin, Michaelian & Sant'Anna present three objections against Fernández' view on the feeling of pastness. First, Fernández pretends to propose a first-order

---

3   Michaelian & Robins' comments on the functionalist theory of memory are based on Fernández (2018), which is a previous version of Chapter 2 of his book.

4   As Sant'Anna & Michaelian indicate, explaining the alleged particularity of remembering is not only a problem for the functionalist theory, but also for the simulation theory of memory.

representationalist account of memory, but the idea that the feeling of pastness is an experience of the causal component of mnemonic content does not seem to be an account of this kind. Second, they suggest that the claim according to which subjects experience some components of mnemonic content is not fully clear: for example, how can something as abstract as the causal component of mnemonic content be an object of introspection? Third, Fernández' proposal is not supported by empirical evidence and, actually, Perrin, Michaelian & Sant'Anna point out that some empirical work demonstrates that the feeling of pastness is sensitive not to the features of the content of memory states, but to the features of the processes that generate those states. With respect to one of the epistemic aspects of remembering, Michaelian asserts that Fernández' thesis that memories are immune to error through misidentification is misguided, because, Michaelian thinks, it is built on both a problematic view of mnemonic content and a questionable definition of observer memory. Contrary to Fernández, Michaelian contends that observer memories do imply that memory is vulnerable to errors through misidentification.

These criticisms highlight that the functionalist theory of memory, the self-referential view of mnemonic content, and the applications of the latter to the phenomenology and the epistemology of memory need further developments and clarifications, to confront important objections, and discuss with rival theories. All in all, this indicates that *Memory: A Self-Referential Account* has already had a considerable impact on the field, and that it will certainly be present in many important debates in the philosophy of memory.

References

Bernecker, S. (2020). Review of the book *Memory: a self-referential account,* by J. Fernández. *Notre Dame Philosophical Reviews. An Electronic Journal.* https://ndpr.nd.edu/news/memory-a-self-referential-account/

Fernández, J. (2006). The intentionality of memory. *Australasian Journal of Philosophy,* 84(1), 39-57. http://dx.doi.org/10.1080/00048400600571695

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, Taylor & Francis Group. https://doi.org/10.4324/9781315159591

Fernández, J. (2019). *Memory: a self-referential account.* Oxford University Press. https://doi.org/10.1093/oso/9780190073008.001.0001

Michaelian, K. (2020). Episodic memory is not immune to error through misidentification: against Fernández. *Synthese.* Advance online publication. https://doi.org/10.1007/s11229-020-02652-w

Michaelian, K. & Robins, S. K. (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, Taylor & Francis Group. https://doi.org/10.4324/9781315159591

Perrin, D., Michaelian, K. & Sant'Anna, A. (2020). The phenomenology of remembering is an epistemic feeling. *Frontiers in Psychology,* 11:1531. https://doi.org/10.3389/fpsyg.2020.01531

Sant'Anna, A. & Michaelian, K. (2019). Teorías sobre o lembrar: causalismo, simulacionismo e funcionalismo. *Voluntas: Revista Internacional de Filosofía,* 10(3), 8-36. https://doi.org/10.5902/2179378640445

Nota a los autores

La revista *Estudios de Filosofía* publica preferentemente contribuciones que provengan de investigaciones en los diferentes campos de la filosofía, así como artículos que aporten a la discusión, el esclarecimiento, la actualización o la interpretación de problemas filosóficos. Se aceptan contribuciones en español e inglés. Es de suma importancia que los artículos enviados cumplan los siguientes requisitos:

1. Entrega

Para realizar envíos de contribuciones a *Estudios de Filosofía* es necesario que los autores estén registrados en el sitio web de la revista (OJS) en el siguiente link: https://revistas.udea.edu.co/index. php/estudios_de_filosofia/about/submissions. Acceda al sitio y proceda al envío, que se hace a través de cinco pasos. Así mismo, lea el aviso de derechos de autor.

Preparación

Todos los manuscritos serán sometidos a doble arbitraje anónimo antes de su posible publicación. Por ello, rogamos a los autores eliminar del cuerpo del texto su nombre, afiliación institucional y cualquier referencia que pueda ayudar a identificarlo.

Puede darse el caso de que al autor se le pida revisar su artículo antes de ser publicado. Además, el derecho de rechazar la publicación de los artículos es reservado al editor. Para información sobre los tipos y extensión de los artículos que la revista publica consulte la tipología de los artículos en la página web de la revista. Sugerimos usar el siguiente instructivo de guía para la preparación de su envío.

Estructura

Proceda según la siguiente estructura para presentar su envío:

- **Título**
- **Resumen:** en inglés y español, entre 130 y 150 palabras en un solo párrafo donde señale los objetivos y propósitos, así como las conclusiones. El resumen no debe contener ninguna abreviatura indefinida ni referencias no especificadas.
- **Palabras clave:** una lista de 5 a 7 descriptores, los cuales son usados con fines de indexación.
- **Texto principal del artículo.**
- **Referencias:** APA séptima edición.

Diseño del manuscrito

El archivo enviado debe estar en .docx.

Página: tamaño carta, todas las páginas numeradas consecutivamente.

Fuente: tamaño 11 o 12 puntos, Times New Roman

Espaciado: 1,5.

Márgenes: 3 cms. arriba, abajo, derecha e izquierda.

2. Datos imprescindibles que deben incluirse en el sistema de envíos (OJS) de la revista

- El artículo debe ir precedido por un resumen de máximo 150 palabras en el que se describan el contenido y la tesis central del artículo, además de 5 a 7 palabras clave para identificar los temas sobre los que versa.
- Al final del artículo debe ir la lista de referencias.
- **En sistema de envíos OJS:** todas las colaboraciones deben aportar una breve biografía intelectual del autor, que no exceda 150 palabras, en la que aparezcan los siguientes datos: nombre y apellidos del autor, filiación institucional completa (universidad, dependencia, grupo de investigación); estudios realizados (títulos obtenidos, institución); libros y artículos publicados recientemente; áreas de especialización; ORCID, e-mail y dirección postal.

3. Formatos para las referencias bibliográficas

La revista se rige únicamente por el modelo de citación APA séptima edición. En el cuerpo del artículo se hace la referencia entre paréntesis indicando el apellido del autor, el año de publicación de la obra y el número de página, p. ej. (Nietzsche, 1973, p. 20).

Cuando se citan obras de un mismo autor publicadas el mismo año, la diferencia se indica mediante una letra minúscula adjunta al dato del año, p. ej. (Gutiérrez, 1999a). Cuando sea necesario repetir referencias a una misma obra, se usan de nuevo los paréntesis para repetir los datos de autor, año o página, según sea necesario.

El formato de referencias bibliográficas, que debe incluirse al final del artículo, se rige por el siguiente orden, puntuación y resalte tipográfico:

- **Libros**

  Nietzsche, F. (1973). *Así habló Zaratustra* (A. Sánchez Pascual, Trad.). Madrid: Alianza.
- **Artículos de revistas**

  García, C. (1985). El problema de la virtud en Platón. *Pensamiento Filosófico, 12*(4), 45-60.
- **Capítulos de libros o colaboraciones en obras colectivas:**

  Gutiérrez, A. (1999). La cuestión del juicio determinante en el pensamiento de Kant. En: M. Arboleda (Ed.)*, El pensamiento de Kant en la reflexión filosófica contemporánea* (pp. 325-350). Madrid: Tecnos.
- **Recursos electrónicos:**

  Pritzker, T. J (s.f.). An early fragment from central Nepal [sitio en internet]. *Ingress Communications,* http://www.ingress.com/~astanart/pritzker.html.

4. Recepción, evaluación y aceptación de contribuciones

- La evaluación de los artículos es anónima. El comité editorial de la revista selecciona a los jurados de un amplio grupo de prestigiosos filósofos colombianos y extranjeros que cubre todas las áreas de la filosofía.
- La revista informará a su debido tiempo los resultados de las evaluaciones de los artículos.
- Con autorización del autor, la revista realizará cambios editoriales cuando sean necesarios.

Las colaboraciones deben ser enviadas a *Estudios de Filosofía* en el siguiente link:
https://revistas.udea.edu.co/index.php/estudios_de_filosofia/about/submissions

Última actualización: 10 de julio de 2020

Código de ética de publicación

1. Introducción: Temática y alcance de *Estudios de Filosofía*

Estudios de Filosofía es la revista editada por el Instituto de Filosofía de la Universidad de Antioquia. Es una publicación electrónica internacional de acceso abierto regida por el sistema de doble arbitraje anónimo. Circula semestralmente de manera ordinaria, sin perjuicio de que, a juicio del Comité editorial, se realicen publicaciones extraordinarias. Desde su fundación en 1990, *Estudios de Filosofía* se ha concebido como medio especializado para el fomento y la difusión de trabajos de investigación en todos los campos de la filosofía, tanto de investigadores colombianos como de miembros de la comunidad filosófica internacional. La institucionalidad de *Estudios de Filosofía* garantiza su orientación hacia el desarrollo de las investigaciones filosóficas en el país y el fortalecimiento de una cultura de comunicación, bajo el principio del respeto a la libertad de expresión e investigación. Se trata de una publicación dirigida a un público de especialistas en filosofía, pero también a todas aquellas personas interesadas en el debate intelectual contemporáneo.

El propósito de publicar este Código de ética de publicación es señalar las expectativas de *Estudios de Filosofía* y el Instituto de Filosofía de la Universidad de Antioquia con respecto a la ética de publicación, teniendo como referente la temática, enfoque y alcance de la revista. Este código presenta los estándares éticos básicos para autores y evaluadores, así como también señala esquemáticamente las funciones y responsabilidades del editor.

2. Estructura editorial

Estudios de Filosofía está coordinada sólo por un director, quien a su vez coordina al editor (general o invitado), al asistente de dirección y al asistente editorial, al diagramador y al equipo editorial. El director es nombrado por el director del Instituto de Filosofía. El director es el principal responsable de la revista y todo lo que esta implica. El editor determina la temática de cada número y evalúa la adecuación temática y la calidad de cada artículo que llega a la revista antes de enviarlo a evaluar. El/la asistente y el comité editorial se encargan de apoyar al editor con la revisión del material potencialmente publicable, así como de la asesoría en el nombramiento de los revisores externos. El diagramador es el responsable del diseño y formato de la revista, y el editor es el encargado de la versión final de cada número publicado de *Estudios de Filosofía*. Todos los cargos antes mencionados son ocupados bajo el principio de la libre voluntad de cada uno de los miembros.

3. Obligaciones y responsabilidades generales del director

El director es responsable de:

3.1. Trabajar por satisfacer las necesidades de los lectores y autores.

3.2. Tratar de mejorar constantemente la revista.

3.3. Tener procesos para asegurar la calidad del material publicado.

3.4. Defender la libertad de expresión.

3.5. El mantenimiento de la integridad del expediente académico.

3.6. Oponerse a las necesidades empresariales que comprometan los estándares intelectuales y éticos de la revista.

3.7. Estar siempre dispuesto a publicar correcciones, aclaraciones, retractaciones y disculpas, cuando sea necesario.

4. Relaciones con los lectores

A los lectores se les

4.1. Informará acerca de las fuentes de financiamiento, las investigaciones y si los financiadores tenían algún papel en la investigación y su publicación y, en caso afirmativo, cuál fue.

4.2. Garantizará que todos los informes publicados y revisiones de la investigación hayan sido revisados por evaluadores calificados.

4.3. Asegurará que las secciones no arbitradas de la revista están claramente identificadas.

4.4. Informará las medidas adoptadas para garantizar que manuscritos de los miembros de la revista o consejo editorial reciban una evaluación objetiva e imparcial.

5. Relaciones con los autores

5.1. Las decisiones del editor de aceptar o rechazar un artículo para su publicación se basarán en la importancia del manuscrito, su originalidad y claridad, así como la validez del estudio y su relevancia para la revista.

5.2. El editor no revocará las decisiones del editor anterior de aceptar manuscritos para su publicación, salvo que se les identifique serios problemas.

5.3. *Estudios de Filosofía* hace una detallada descripción de los procesos de revisión por pares, y el editor justificará cualquier desviación importante de los procesos descritos.

5.4. *Estudios de Filosofía* tiene un mecanismo para que los autores apelen las decisiones editoriales, a través de la comunicación con el Comité editorial.

5.5. *Estudios de Filosofía* ha publicado una guía para los autores en la que se señala qué espera de sus manuscritos. Esta guía regularmente se actualiza y señala un vínculo para acceder a esta sección.

6. Compromisos del editor

El editor de *Estudios de Filosofía* se compromete a:

6.1. Orientar a los evaluadores en el proceso de evaluación, incluso en la necesidad de manejar el material evaluado con confidencialidad.

6.2. Exigir a los evaluadores que declaren los posibles conflictos de intereses antes de aceptar evaluar un manuscrito.

6.3. Contar con sistemas adecuados para asegurar que las identidades de los evaluadores estén protegidas.

6.4. Impulsar a que los revisores que comenten asuntos éticos y de la investigación, así como la posible mala conducta de publicación planteada en los manuscritos.

6.5. Estimular a los revisores para que comenten la originalidad de los manuscritos y a que estén atentos a si la publicación es redundante o constituye plagio.

6.6. Enviar, en su totalidad, los comentarios de los evaluadores a los autores, salvo que contengan términos injuriosos o difamatorios.

6.7. Reconocer la contribución de los evaluadores a la revista.

6.8. Monitorear el desempeño de los evaluadores y tomar medidas para garantizar que sean de alto nivel.

6.9. Desarrollar y mantener una base de datos de evaluadores idóneos y actualizarla con base en el desempeño del evaluador.

6.10. Dejar de contactar a los evaluadores que de manera recurrente hacen evaluaciones deficientes, descorteses o de baja calidad.

6.11. Utilizar una amplia gama de fuentes (no sólo los contactos personales) para identificar posibles nuevos evaluadores (por ejemplo, bases de datos bibliográficas).

7. Relaciones del editor con el Comité editorial

El editor de *Estudios de Filosofía* proporciona a los nuevos miembros del Comité editorial las directrices sobre todo lo que se espera de ellos y mantiene actualizados sobre las nuevas políticas y desarrollos a los miembros existentes. Además de esto el editor:

7.1. Tiene políticas establecidas para el manejo de los manuscritos de los miembros del Comité editorial para asegurarles una revisión imparcial.

7.2. Identifica miembros para comité editorial, debidamente cualificados, que puedan contribuir activamente al desarrollo y buena gestión de la revista.

7.3. Examina regularmente la composición del comité editorial.

7.4. Ofrece una orientación clara a los miembros del comité editorial acerca de sus funciones y deberes previstos, que incluyen: (1) actuar como embajadores de la revista, (2) apoyar y promover la revista, (3) buscar los mejores autores y trabajos, y fomentar activamente el envío de manuscritos, (4) revisar los envíos a la revista, (5) aceptar ser comisionados para escribir editoriales, críticas y comentarios sobre artículos en su área de especialización, (6) asistir y contribuir a las reuniones del comité editorial.

7.5. Consulta periódicamente a los miembros del comité editorial para conocer sus opiniones sobre la marcha de la revista, informarles cualquier cambio en sus políticas, e identificar los retos del futuro.

8. Procesos editoriales y de evaluación por pares

El editor de *Estudios de Filosofía* vela por:

8.1. Esforzarse por garantizar que la revisión por pares en la revista es justa, imparcial y oportuna.

8.2. Tener sistemas para asegurar que el material remitido para su publicación es confidencial durante el proceso de evaluación.

8.3. Toma todas las medidas razonables para asegurar la calidad del material publicado, reconociendo que las revistas y secciones dentro de las revistas tienen objetivos y normas diferentes.

9. Manejo de casos en que posiblemente se incurra mala conducta

9.1. El editor de *Estudios de Filosofía* tiene el deber de actuar cuando haya sospecha de mala conducta o cuando esta sea alegada. Esta obligación se extiende a los artículos publicados y no publicados.

9.2. El editor no rechazará simplemente los documentos que plantean dudas sobre posible mala conducta, está éticamente obligado a investigarlos.

9.3. El editor sigue los diagramas de flujo que el COPE sugiere para estos casos.

9.4. El editor solicitará primero una respuesta de quienes se sospeche que incurren en mala conducta. Si no queda satisfecho con la respuesta, debe pedir a los empleadores respectivos, a la institución, o alguna instancia apropiada, como de la organización nacional de investigación integridad, que se adelante la respectiva investigación.

9.5. El editor deberá hacer todos los esfuerzos razonables para asegurar que la debida investigación de supuesta mala conducta se lleva a cabo. Si esto no ocurre, deberá hacer todos los esfuerzos razonables para persistir en la obtención de una solución al problema. En *Estudios de Filosofía* somos conscientes de que esto es un deber dispendioso, pero importante.

10. Fomento del debate académico

10.1. El editor promueve y está dispuesto a considerar para su publicación las críticas académicas a los trabajos publicados en esta revista.

10.2. Los autores de los trabajos criticados tienen la oportunidad de responder.

10.3. El editor está abierto a la investigación que cuestiona el trabajo anterior publicado en la revista.

11. Conflictos de intereses

11.1. El editor tiene formas para gestionar sus propios conflictos de intereses, así como los de autores, revisores y miembros del comité editorial.

11.2. *Estudios de Filosofía* tiene un proceso específico para el manejo de los manuscritos de los editores, profesores del Instituto de Filosofía o miembros del consejo del comité académico para asegurar una revisión imparcial.

Este código se realizó usando como base las guías que ofrece el Committee of Publication Ethics (COPE).

Última actualización: 10 de julio de 2020.

Ethics guidelines

1. Introduction: focus and scope of *Estudios de Filosofía*

Estudios de Filosofía is the journal published by the Institute of Philosophy of the Universidad de Antioquia. It is an international peer-reviewed, open-access, electronic journal and adheres to the policy of double-blind peer review. Since its foundation in 1990, *Estudios de Filosofía* has been devoted to fostering the research in all fields of philosophy. The journal publishes papers in Spanish and English.

The purpose of publishing these code of ethics is to indicate the expectations of the journal and of the Institute of Philosophy regarding the publishing ethics, having as a referent the theme, focus and scope of the journal. This code presents basic ethical standards for authors and reviewers, and also points out schematically the functions and responsibilities of the editor.

2. Editorial Structure

Estudios de Filosofía is coordinated by one director, who at the same time coordinates the (general or guest) editor, the editorial assistant, the designer and the editorial team. The director is appointed by the director of the Institute of Philosophy. The director is the main responsible for the journal. The editor determines the theme of each issue and evaluates the thematic relevance and quality of each article submitted to the journal before sending it to review. The editorial assistant and the editorial committee support the editor by reviewing potentially publishable material, as well as by suggesting external reviewers. The designer is responsible for the style and format of the journal. Finally, the editor is in charge of the final version of each published issue of *Estudios de Filosofía.*

3. General Obligations and Responsibilities of the Director

The director is responsible for
3.1. Working to satisfy the necessities of both readers and authors.
3.2. Trying to improve constantly the journal.
3.3. Put in place processes to ensure the quality of the published material.
3.4. Defending free speech.
3.5. The integrity of the academic archive.
3.6. Opposing to mercantile tendencies that may compromise the intellectual and ethical standards of the journal.
3.7. Being always willing to publish corrections, clarifications, retractions and apologies when necessary.

4. Relations with the readers.

Readers will be

4.1. Informed about the financial sources of the researches, whether the financers had any role in the research and its publication and, if so, which one.

4.2. Guaranteed to have all the published reports and research revisions checked by qualified reviewers.

4.3. Assured that the non-reviewed sections of the journal are clearly identified.

4.4. Informed about the measures adopted to guarantee that the manuscripts by the members of the editorial committee receive an objective and unbiased evaluation.

5. Relations with the Authors

5.1. The editor's decision concerning the acceptance or rejection of an article for its publication will be based on the quality of the manuscript, its originality and clarity, as well as its relevance for the journal.

5.2. The editor will not revoke the former editor's decision to accept manuscripts for their publication, unless serious problems were identified.

5.3. *Estudios de Filosofía* makes a detailed description of the peer-review processes, and the editor will justify any important deviation from the described processes.

5.4. *Estudios de Filosofía* has a mechanism for authors to appeal the editorial decisions, through communication with the editorial committee.

5.5. *Estudios de Filosofía* published guidelines for the authors where they can check what is expected from their manuscripts. These guidelines are regularly updated.

6. Commitments of the Editor

The editor commits to:

6.1. Guiding the review process in order to guarantee the quality and integrity of it. For this purpose, the editor designed guidelines and a review format which will be periodically updated.

6.2. Requiring the reviewers to declare the possible conflicts of interests before agreeing to review a manuscript.

6.3. Having adequate systems to ensure the protection of the reviewer's identities.

6.4. Encouraging the reviewers to comment on ethical and research issues, as well as possible misconducts.

6.5. Prompting the reviewers to comment on the originality of the manuscripts and to be vigilant about redundancies or plagiarism in the publication.

6.6. Sending to the authors the totality of the reviewer's comments, unless they are injurious or defamatory.

6.7. Acknowledging the reviewer's contribution to the journal.

6.8. Monitoring the reviewer's performance in order to guarantee the quality of the review.

6.9. Developing and keeping a suitable reviewer database and updating it based on the reviewer's performance.

6.10. Avoiding the reviewers who regularly make poor, unkind or low-quality reviews.

6.11. Using a wide range of sources (not only personal contacts) to identify possible new reviewers (e.g. bibliographical databases).

7. Relations of the Editor with the Editorial Committee

The editor of *Estudios de Filosofía* provides the new members of the editorial committee with guidelines on everything expected from them and shares updates about new policies and developments with the existing members. In addition, the editor:

7.1 Has established policies for managing the manuscripts of the editorial committee members in order to ensure an unbiased review.

7.2 Identifies qualified members for the editorial committee who can actively contribute to the development and good management of the journal.

7.3 Regularly examines the conformation of the editorial committee.

7.4 Offers the members of the editorial committee a clear orientation about their expected duties and functions, including: (1) to act as ambassadors for the journal, (2) to support and promote the journal, (3) to search for the best authors and works, and actively encourage the submission of manuscripts, (4) to accept being commissioned to write editorials, critiques and commentaries on articles in their areas of expertise, (5) to assist and contribute to the editorial committee meetings.

7.5 Consults periodically the members of the editorial committee to know their opinions on the performance of the journal. The editor also informs them about any change on its policies and identifies challenges for the future.

8. Editorial and Peer-Review Processes

The editor of *Estudios de Filosofía* watches over

8.1 Making an effort to guarantee a fair, unbiased and timely peer review.

8.2 Having systems to ensure that the submitted material for publication is confidential during the review process.

8.3 Taking reasonable measures to ensure the quality of the published material, recognizing that the sections within the journals have different goals and norms.

9. Managing Cases of Possible Misconduct

9.1 The editor of *Estudios de Filosofía* has the duty to act whenever there is suspicion or report of misconduct. This applies to both published and unpublished articles.

9.2 The editor will not simply reject the documents that raise questions about possible misconduct, s/he is ethically obligated to investigate them.

9.3 The editor will follow the COPE guidelines for these cases.

9.4 The editor will first request an answer from those under suspicion of misconduct. If the answer is not satisfactory, an investigation must be solicited to the employers, institution, or the research agency.

9.5 The editor shall make all the reasonable efforts to ensure that a proper investigation of alleged misconduct is carried out. If this does not happen, the editor must make all the reasonable efforts to persist in obtaining a solution to the problem. In *Estudios de Filosofía* we are aware of the difficulty of this duty, but also of its importance.

10. Encouragement of Academic Debate

10.1 The editor promotes and is willing to consider for publication academic critiques of the works published in this journal.

10.2 The authors of the critiqued works will have the opportunity to reply.

10.3 The journal accepts papers that criticize previous publications of the journal.

11. Conflict of Interest

11.1 The editor has ways to manage his/her own conflicts of interest, as well as those of the authors, reviewers and members of the editorial committee.

11.2 *Estudios de Filosofía* has a specific process for managing the manuscripts by the editors, members of the Institute of Philosophy (UdeA) or members from the editorial committee, to ensure an unbiased review.

This code is based on the guidelines provided by the Committee of Publication Ethics (COPE).

(Last update: October 8, 2020)