Cryptomnesia: a three-factor account

Christopher Jude McCarroll[1] · André Sant'Anna[2]

Abstract

Understood as a psychological phenomenon, there has been very little discussion of cryptomnesia in the philosophical literature. Cryptomnesia presents us with a strange phenomenon in which we take ourselves to be imagining, but the thought or idea that we entertain actually involves remembered content. In this paper, we argue for a three-factor account of cryptomnesia, according to which it is a mnemonic phenomenon that involves imagination. We provide an account of both the 'mnemonic' and 'imaginative' aspects of cryptomnesia in terms of the attitude, the content, and the metacognitive processes involved in those states. In addition, we show how our three-factor account is better suited to account for cryptomnesia than competing philosophical theories of episodic memory. We conclude by discussing how the three-factor account sheds light on a range of other mnemonic and imaginative phenomena.

> You don't set out to try and steal anything, but it can happen…If you write all the time, you're going to collect those things and not know it.

—Dolly Parton[1]

---

[1] This quotation is from an interview Dolly Parton conducted with the BBC; see Savage (2016).

Christopher Jude McCarroll and André Sant'Anna have contributed equally to this paper.

✉ Christopher Jude McCarroll
  chrismccarroll@nycu.edu.tw

[1] Institute of Philosophy of Mind and Cognition, National Yang Ming Chiao Tung University, No. 155, Sec. 2, Linong Street, Taipei 112, Taiwan, ROC

[2] Cologne Center for Contemporary Epistemology and the Kantian Tradition (CONCEPT), University of Cologne, Albertus-Magnus-Platz, 50923 Cologne, Germany

1 Introduction

Personal memory permeates many aspects of human life.[2] From the judicial system, to personal relations and moral concerns, through to one's sense of diachronic identity, memory is central to many aspects of what it means to live a rich and meaningful life (e.g., Craver, 2020; Gerrans & Kennett, 2017; Schechtman, 2001). Given the important role memory plays in these diverse facets of human life, it is important to know the difference between remembering and other similar states. It is often vital to be able to distinguish genuine memories from mere imaginings that are mistaken for memories. A lot can hang on whether one's memory is genuine or not. It is crucial to distinguish, for example, whether one did in fact perceive a crime and hence episodically remembers it, or whether one merely imagined, dreamed, or heard about the event, and then mistakenly thinks that one is remembering a real event that one witnessed in the past.

Yet it can be just as important to identify what may in fact be remembered content, but is taken to be imagined. Cryptomnesia refers to precisely this phenomenon (Brown & Murphy, 1989; Marsh et al., 1997). In cases of cryptomnesia, one takes oneself to be imagining, but the thought or idea that one entertains actually involves remembered content.[3] Cryptomnesia can often result in unconsciousness plagiarism or even self-plagiarism (Skinner, 1983). You remember an image, a song, or an idea that you previously experienced, but you mistake this for something you created or invented. In such cases you appear to be unsuccessfully imagining.

Given the social nature of many creative endeavours, such as science and musical creation, there are numerous anecdotal, and controversial, examples of cryptomnesia. Carl Jung suggested that Nietzsche, writing in *Thus Spoke Zarathustra*, included an almost exact account of an incident that appeared in a book published 50 years before Nietzsche's writings. Rather than accuse Nietzsche of deliberate plagiarism, Jung believed that he had cryptomnesia, or the 'concealed recollection of a textual memory' (Oppel, 2005, p. 204). According to Jung, the previous text had 'secretly crept up and reproduced itself' in Nietzsche, and 'shows how the unconscious layers of the mind work' (Jung, quoted in Oppel, 2005, p. 204).

The world of musical creation, too, is littered with examples of cryptomnesia. One of the most famous legal cases of copyright infringement involved George Harrison, of the Beatles. Harrison's song, *My Sweet Lord*, seemed to bear an uncanny resemblance to an earlier song composed by Ronald Mack for the Chiffons, entitled *He's So Fine*. The resemblance resulted in a lawsuit that was brought against Harrison. During the trial, Harrison admitted that he had heard the Chiffon's song, but denied that he intentionally copied it. Even though it was found that Harrison did infringe copyright, because of the similarity between the songs and the fact that Harrison had accessed the earlier one, he was not accused of deliberate plagiarism: 'Did Harrison deliberately use the music of "He's so Fine?" I do not believe he did,' explained the judge who presided

---

[2] In this paper we are concerned with what's known as "personal" or "episodic" memory (we treat these terms as synonymous). Roughly, this is the type of memory that allows one to recall events in one's personal past (Sutton, 1998).

[3] We go on to clarify this claim in Sect. 5. Our claim will be that cryptomnesia is not an instance of remembering even though it is a mnemonic phenomenon.

over the case (Yin, 2016).[4] The line between deliberate plagiarism and cryptomnesia may be fine, but it is an important one, and one that a theory of remembering needs to account for.

How can we answer such dilemmas about authentic creation? Is cryptomnesia remembering? Or is it imagining? In this paper, we argue for a three-factor account of cryptomnesia, according to which it is a mnemonic phenomenon that involves imagination. While this may sound counterintuitive on the face of it, we will show that this seeming counterintuitiveness is in fact illusory, and that understanding cryptomnesia in this way helps elucidate its important features.

We proceed in the following way. We first provide a more detailed characterization of cryptomnesia and propose a discussion of it in relation to recent theories of episodic memory (Sect. 2). Next, we consider how the causal theory of memory (Sect. 3) and the simulation theory (Sect. 4) might account for cryptomnesia. We find the casual theory's account to be wrong, and we find the simulation theory's account to be incomplete. We then outline our own three-factor account of cryptomnesia (Sect. 5), which explains both the mnemonic and imaginative dimensions of the phenomenon. Finally, in Sect. 6, we differentiate our account from those that feature an epistemic relevance condition.

2 Characterizing cryptomnesia

Cryptomnesia is a mnemonic phenomenon that involves imagination. In this section, we provide a more detailed discussion of what exactly is 'mnemonic' and what is 'imaginative' about cryptomnesia. We also discuss why an account of cryptomnesia can be plausibly grounded on an account of episodic memory.

Cryptomnesia presents us with a strange phenomenon, in the sense that, subjectively speaking, it seems to be a creative or imaginative act. However, cognitively speaking, it also seems to involve the retrieval of retained information originating in past experience, which the subject fails to identify as such. This brief characterization highlights four important features that we consider to be defining of cryptomnesia. First, cryptomnesia seems to involve a *phenomenology* that is characteristic of at least some types of imaginative states. As the examples discussed above illustrate, cryptomnesic subjects clearly seem to have an experience of imagining the relevant events or objects. We will call this feature PHENOMENOLOGY.

Note that in highlighting PHENOMENOLOGY as a feature of cryptomnesia that needs to be explained, we should not be read as claiming that making sense of this particular feature requires providing an account of the *phenomenal character* of, or *what it is like* to have, cryptomnesia. The task of providing an account of the phenomenology of imagination is a notoriously difficult one (see Kind, 2020), and it is

---

[4] There is another, controversial and disputed, case of cryptomnesia involving the Beatles. According to Sir Paul McCartney, the song *Yesterday* came to him in his sleep; it was something he must have dreamed—it "fell out of bed". The musicologist Spencer Leigh, however, claims that the song *Yesterday* was "subconsciously influenced" by a Nat King Cole song called *Answer me*. Leigh does not suggest that the former Beatle plagiarised *Answer me*, but the thought is that McCartney had previously heard, and was influenced by, the song, and then *Yesterday* appeared to him as springing from his own creative imagination (see Womack, 2016, p. 559).

one that we do not intend to take on here. Moreover, as many philosophers of imagination have noted, there are different ways in which a mental state can be said to be imaginative (Kind, 2016; Langland-Hassan, 2021; Van Leeuwen, 2013, 2014), which suggests that there are equally different ways in which one may have an experience of imagining.[5] Thus, our claim is that a satisfactory account of PHENOMENOLOGY only requires an explanation *why* cryptomnesia is experienced as being imaginative, but not (or at least not necessarily) of *what kind* of imaginative experience it is. In other words, PHENOMENOLOGY only requires an explanation of why cryptomnesia is experienced as being imaginative as opposed to some other type of state.

Second, cryptomnesia seems to be a *creative* act, in the sense that subjects entertain the content of those states as possible or novel, such that they can be changed or altered by them as they see fit.[6] We will call this feature CREATIVITY. Although CREATIVITY and PHENOMENOLOGY are related in important ways—e.g., engaging in a creative act seems to presuppose that a subject experiences a mental state as an imagining—there's also a sense in which they are importantly different. Consider cases in which subjects imagine events that actually took place in the past, such as events associated with the Waterloo Battle. Those states are arguably experienced as cases of imagining, but they are not necessarily experienced as creative. A subject can, for instance, imagine those events after reading about them in a history book, and as long as the subject is trying to remain accurate to the actual happenings, this imagining does not exhibit CREATIVITY. This doesn't mean, of course, that it couldn't exhibit CREATIVITY—e.g., the subject might imagine how the Waterloo Battle would have turned out if some events had unfolded in a different way—but only that experiencing a state as an imagining does not necessarily mean that one experiences it as a creative act.[7]

One important qualification about CREATIVITY is that the contents of cryptomnesic states need not actually be creations of the subject. All that is required by CREATIVITY is that they are *experienced* as being the result of a creative act. Thus, for instance, one might be in a cryptomnesic state even if the content of one's mental state is identical to a content experienced—e.g., even if a song that one takes to be a result of one's mental state is a note-by-note reproduction of a song created by someone else. What matters for cryptomnesia is that one experiences that content as being the result of one's own creative act. This allows for a distinction between 'preserved' and 'modified' cases of cryptomnesia, with *preserved cryptomnesia* referring to cases where the content entertained is identical (or very similar) to the content that was originally experienced, and *modified cryptomnesia* referring to cases where the content entertained has actually been altered by the subject (see Sect. 5).

Third, cryptomnesia seems to involve the retention of information acquired in the past, such that this information is brought to consciousness in the present and becomes the object of a creative act. Consider the Harrison example discussed above. The fact that Harrison was found to have infringed copyright vindicates our intuition that at least

---

[5] We return to this issue in our discussion of our positive proposal in Sect. 5.

[6] This is an importantly different sense of creativity than that invoked in the literature on constructive memory, where memory representations are often said to be creatively constructed by using various sources of information. The subject is not (typically) in control of this creative construction, however.

[7] See Sect. 5 for a more detailed discussion of this point.

part of the information that figured in *My Sweet Lord* was drawn from one of Harrison's previous encounters with Mack's song. We will call this feature RETENTION.

Finally, the fourth feature is that cryptomnesia appears to involve a failure by the subject to identify the content of their mental state as originating in an experience previously had by them. Again, as the Harrison case nicely illustrates, Harrison doesn't deny having heard Mack's song before; what he denies is that his having heard the song before played a role in his creation of *My Sweet Lord.* We will call FAILURE the subject's failure to identify the source of the information contained in their mental state in cases of cryptomnesia. Both RETENTION and FAILURE are what we refer to as the 'mnemonic' dimension of cryptomnesia.

In what follows, we will discuss how two prominent philosophical theories of episodic remembering—the causal theory and the simulation theory—fare in relation to these four features. Before we move on to this discussion, we pause to address two potential concerns that readers might have in connection to our focus on these theories. The first concern has to do with RETENTION. In particular, one might worry that RETENTION is a theoretically-laden notion, favouring a causalist approach to remembering, where successful remembering requires, among other things, the existence of memory traces that link past experiences to memories (see Robins, 2016; Michaelian & Robins, 2018). RETENTION would not, therefore, be compatible with the simulation theory, which explicitly denies that memory traces are necessary for successful remembering (Michaelian, 2016a; see Sect. 4). But if that is the case, then our characterization of cryptomnesia puts simulationist approaches at a significant disadvantage.

In response, we clarify that our use of RETENTION is intended to be compatible with multiple theories of remembering, including causal and simulation theories. On our usage of the term, information is retained whenever a subject draws on a memory trace that carries information about the event in question. We remain silent, however, on whether such a trace only stores information originating exclusively in a single event.[8] This way of understanding memory traces is compatible with both the causal and simulation theories of remembering, because as Michaelian himself notes 'the simulation theory does not deny that remembering draws on traces but only that a memory of a given event must draw on a trace deriving from the subject's experience of that event' (Michaelian 2023, pp. 83–84).

These considerations point to a second concern one might have with our approach. Both the causal theory and the simulation theory are theories of *episodic* remembering—i.e., occurrences of remembering that are about specific events in our personal past. But even if we assume that there is something mnemonic about cryptomnesia, it is unclear why we should approach this mnemonic dimension of the phenomenon in relation to episodic memory. Couldn't it be that this aspect is explained in terms of the relationship between cryptomnesia and other forms of memory, such as semantic memory? While addressing this question is beyond the scope of this paper, we think that there is an important reason for thinking that the mnemonic dimension of cryptomnesia cannot be fully explained in terms of semantic memory.

---

[8] Traces on this way of understanding them might be more like props (Langland-Hassan, 2022a, b) or even involve semantic information. For other accounts of memory traces, see, for example, Robins (2016) and De Brigard (2014).

This has to do with FAILURE above. What seems to be puzzling about cases of cryptomnesia is not only the fact that there is retention of information, but also that we fail to recognize the entertained content as *originating* in our past experiences. This becomes clearer by looking again at the Harrison example. As we noted before, Harrison acknowledges having heard Mack's song before—which, we suggest, indicates that he semantically remembers it—but he denies that the content of his song originates in one of the experiences he (Harrison) had of the song. Thus, if the mnemonic character of cryptomnesia were simply restricted to the operation of semantic memory, it would be hard to make sense of why Harrison was not charged with deliberate plagiarism, for on this account, he would count as genuinely remembering the song.[9]

To this, one may respond by saying that an equally plausible way of explaining why Harrison was not charged with deliberate plagiarism is because he failed to form the *belief* that he was remembering the song. This would be compatible with Harrison remembering the song semantically, and consequently at odds with our suggestion to treat cryptomnesia as episodic in nature. However, we don't think that's right. Perhaps the fact that Harrison failed to form a belief to this effect would be enough to explain why, from a legal point of view, he was not charged with deliberate plagiarism. But this fails to explain the more basic intuition that Harrison's failing has to do with a failure of *memory*. If all there is to the cryptomnesia is semantic remembering, then *nothing* has gone wrong with Harrison's memory. His failure would be, on this view, rather *epistemic* in nature. But cryptomnesia doesn't—or at least it doesn't as a matter of necessity—involve an epistemic failure.

Another related worry concerning our focus on episodic memory is that cryptomnesia does not seem to require a representation with 'episodic' content. It might be argued that, at least sometimes, all there is to cryptomnesia is the retrieval of 'semantic' content, such as when one entertains a sentence from a book in mind and takes that content to be a creation of one's own mind. In response, whether this poses a problem to our view will depend on what one means by 'episodic' and 'semantic' content. If by semantic content it is meant purely linguistic content expressed in the form of a proposition, and if by episodic content it is meant purely imagistic content, then it is clearly the case that cryptomnesia will sometimes—and arguably many times—involve semantic content only. However, we are not committed to such a view of episodic and semantic memory contents. Episodic memory, on our view, can happen even when all one can represent about an event is having encountered something before, such as a sentence. For instance, one can episodically remember having encountered the sentence 'It's like déjà vu all over again' before, even when one cannot recall when, where, or under what conditions such an encounter took place. As long as one entertains the relevant content as originating in the past, that suffices for that content to count as episodic. Such a view is not, of course, uncontroversial, but it is a plausible one and one that has been defended in recent discussions.[10] More importantly, if this is correct, then

---

[9] We're not trying to make sense of legal practices around cryptomnesia, but only the intuitions that drive, at least in part, those practices (i.e., it is because we think the subject is remembering that we think it is justified to legally punish him).

[10] See, e.g., Fernández (2019), for an extensive defence of a view along these lines. See also Dokic (2014, 2021) and Perrin et al. (2020) for sympathetic approaches. We note, moreover, that characterizing episodic memory in terms of an experience of origin is not incompatible with autonoesis. As Fernández (2019)

treating cryptomnesia as an episodic memory-related phenomenon gives us a clear and straightforward answer to the question of why cryptomnesia involves a *memory error*: i.e., because we fail to represent the relevant content as originating in the past.

Perhaps it could further be argued that cryptomnesia may involve bringing back to mind content that was experienced repeatedly in the past, and so cryptomnesia may again be a semantic phenomenon. For example, one's memory of the line, 'All the world's a stage, and all the men and women merely players', is semantic in that it does not derive from a particular experience one had encountering the line. One simply knows this quotation, without recalling any episodes of encountering it. Now, if one can be cryptomnesic with respect to remembered quotations, or songs, or other contents one has encountered many times, and entertainment of which does not bring to mind any details of any particular encounters, then cryptomnesia is not a purely episodic phenomenon.[11] It is important to note, however, that the idea that episodic memory is only about single events in the past is *not* well-supported by empirical evidence. Episodic memory may frequently be about memories of repeated events (Andonovski, 2020). Further, if cryptomnesia were a problem of semantic memory, then, as we suggested above, there is a puzzle about why the subject fails to know that the content of their mental state derives from the past. Semantic memory typically does not involve source monitoring errors, but this is precisely part of the error involved in cryptomnesia.

On a more practical note, our focus on episodic memory is also partly motivated by the fact that the only discussion of cryptomnesia in the philosophical literature that we are aware of—the one provided by Michaelian (2016a, pp. 173–174)—takes place in the context of the simulation theory, which is a theory of episodic memory. Moreover, as we discuss in more detail in Sect. 3, a central case for recent debates in the philosophy of memory—the *painter case* (Martin & Deutscher, 1966)—which many philosophers judge to be a case of episodic remembering, is in fact structurally analogous to cases of cryptomnesia. The frameworks offered by philosophical theories of episodic remembering thus seem to offer a natural starting point for a more thorough investigation of cryptomnesia. Moreover, as we show below, both the classical causal theory and the simulation theory classify cryptomnesia as episodic remembering (in the current literature), whereas we provide an account of why it is *not* episodic remembering, so even if cryptomnesia turns out not to be an episodic memory phenomenon (but just a declarative memory phenomenon) our account is better placed to explain it.

With these clarifications in mind, we now proceed to evaluate whether the causal theory and the simulation theory can explain cryptomnesia. We argue that both face important difficulties because they focus on the *objective* conditions for successful remembering, which involve considerations that take a third-personal or system-level perspective (such as how the content of an apparent memory is produced and where it comes from) at the expense of subjective conditions (what the person takes themselves to be doing). Both theories prioritise the backward-looking casual role of remembering,

---

Footnote 10 continued

argues, the experience of time and self in episodic memory can be explained by the fact that episodic memory represents itself as originating in the past.

[11] We thank an anonymous reviewer for posing this interesting objection.

the way it was produced, and fail to properly consider its forward-looking causal role, the way in which it might lead to certain beliefs, etc. (Langland-Hassan, 2022a, b).

3 The causal theory and cryptomnesia

How would causal theorists classify cases of cryptomnesia? According to the causal theory (Martin & Deutscher, 1966), for a subject to count as remembering an event, she must have experienced the event in the past, and currently represent this event more or less accurately. In addition, these two representations need to be appropriately causally connected, which means that the current representation is caused by a memory trace formed on the basis of a past experience. Together, these conditions provide us with the full causalist analysis of remembering.

When we look at cases of cryptomnesia it becomes clear that they satisfy the three conditions postulated by the causal theory, and hence cryptomnesic subjects are episodically remembering. That is, cryptomnesic subjects represent an event or object now, they had an experience of that event or object before, and the representation they form now is accurate (within certain limits) with respect to the past event or object.

To see better why cryptomnesia is remembering on the causal theory, consider the notorious painter example (Martin & Deutscher, 1966). Suppose that a painter is asked to paint an imagined scene, which leads him to paint a farmyard scene that he genuinely believes to be imagined. However, when the painter's parents see the painting, they recognize it as being a scene once witnessed by the painter as a child. Should we count the painter's mental state as an occurrence of remembering or mere imagining? Martin and Deutscher's (1966) suggestion is that if there is a memory trace formed on the basis of a prior experience causing the painter's mental state in the present, then the painter counts as remembering even if he takes himself to be imagining the scene. Thus, the painter case shows that, for the causal theory of memory, subjective or first-personal criteria are not relevant for making ascriptions of successful remembering.[12] The causal theory's account of remembering and related phenomena primarily involves consideration of objective factors, whereas what the subject takes themselves to be doing is not important. The painter case, which is structurally similar to cases of cryptomnesia, comes out as a case of remembering on the causal theory.[13]

---

[12] There are, of course, different versions of the causal theory, and the question that now arises is if all versions of the theory would classify cryptomnesia as episodic remembering. While it is beyond the scope of this paper to explore this issue in detail, it seems relatively clear that some versions (would) follow the classical view in describing cryptomnesia as remembering. Bernecker (2010), for example, discusses cases of 'ignorant remembering', which can be considered cryptomnesic cases and are classified as remembering. It seems that Werning's (2020) trace minimalist approach would also classify the phenomenon as remembering provided that a minimal trace connects the current experience with the past event. In contrast, it seems that other versions of the causal theory (e.g., Debus, 2010; Fernández, 2019; Robins, 2020a) would reject the idea that cryptomnesia is remembering. We return to these latter versions of the causal theory in the last section and distinguish our view from them.

[13] Although the painter case is not our focus here, it seems to exhibit all the four characteristics of cryptomnesia discussed above, which suggests that it is, at least at the level of the psychological processes involved, a case cryptomnesia (or, perhaps more aptly, a case of 'self-cryptomnesia'). However, since the painter is

The causal theory hence only provides a partial explanation of the mnemonic dimension of cryptomnesia. It can make sense of RETENTION, but since subjective conditions have no space in the causalist story, FAILURE is left unexplained. The causal theory also leaves unexplained the imaginative dimension of cryptomnesia, i.e., PHENOMENOLOGY and CREATIVITY. In this sense, the causal theory's account of cryptomnesia is unsatisfactory, for it fails to account for the personal level or subjective dimension of the phenomenon. It treats cryptomnesia as an instance of successful remembering rather than as involving an error.

Importantly, the causal theory does not provide an account of what distinguishes cases of successful remembering that are cryptomnesia from those that are not. It simply lumps cryptomnesia, in which the subject takes himself to be imagining, in with ordinary cases of remembering (where the subject takes himself to be remembering), but no account is given of the difference between these two. Of course, the causal theory could bring in a metacognitive monitoring element (see Sects. 4 and 5), which would help explain features such as FAILURE and, perhaps, some aspects of PHENOMENOLOGY. Nonetheless, this would still leave CREATIVITY unexplained. The full phenomenology of cryptomnesia, which includes the experience of creativity, cannot be explained by positing a source monitoring failure. As Brown and Murphy explain,

> Source amnesia is similar to cryptomnesia in that information concerning the "context" in which the fact was first experienced has been forgotten. However, with cryptomnesia the recalled information is perceived as original, whereas with source amnesia it is not. (1989, p. 432)

We return to the causal theory and problems with its take on cryptomnesia in Sects. 5 and 6. For now, we turn to the causal theory's main rival.

4 The simulation theory and cryptomnesia

On the simulation theory, remembering is a form of imagining. How would simulation theorists classify cases of cryptomnesia? It might initially be thought that the problem of cryptomnesia does not apply to the simulation theory. If remembering is merely one form of imagining, then there is a sense in which, in cryptomnesia, the subject is necessarily engaged in imagining on some level. The idea would be that because successful remembering, other imaginative simulations, and cryptomnesic states are all products of the same episodic construction system, then we don't need to distinguish between them in the first place. All these states can be classified as states of the same kind, i.e., imaginings. Yet even on the simulation theory, cryptomnesia is a phenomenon

---

Footnote 13 continued

not, strictly speaking, copying someone else's work, but rather using information acquired in a previous experience had by him, it is unclear whether it should ultimately be classified as a case of cryptomnesia. This points to an interesting question for future work, which is whether there is a *normative* dimension to the phenomenon of cryptomnesia. In other words, it might be that in addition to the relevant psychological processes that we discuss here, what we call 'cryptomnesia' is defined in part by how those processes are related to certain social practices that involve norms, such as attributions of intellectual creation. This would help explain why, at least intuitively, we don't classify the painter case as a case of cryptomnesia, for the content of his mental state is not derived from someone else's creation.

that needs to be explained. We need to be able to distinguish which form of imaginative process we are engaged in: 'we still need to explain how subjects distinguish between the form of imagination dedicated to reconstructing past episodes—that is, remembering—and other forms of episodic imagination' (Michaelian, 2016a, p. 175). Cryptomnesia requires a simulationist explanation.

The simulation theory sets two conditions for successful remembering: a subject must represent an event or object in the present and the representation must be produced by a reliably functioning episodic construction system that aims to represent an event from the subject's personal past. Cases of cryptomnesia involve the representation of an event or object in the present. The crucial question for the simulationist is, then, whether that representation is produced by a reliably functioning episodic construction system that aims to represent an event from the subject's personal past.

Let us begin by considering the reliability point. Is the episodic construction system functioning reliably in cryptomnesia? While settling this question cannot be done on conceptual grounds, there does not seem to be any strong reasons to think that cryptomnesia involves any malfunctioning. Cryptomnesia, as Michaelian himself acknowledges, involves an error, but the fact that it is erroneous does not mean that the episodic construction system is working unreliably.[14] So, in the absence of any reasons to the contrary, we take it to be a plausible assumption that cryptomnesic states could be produced by a reliably functioning episodic construction system.

If the process is reliable, then the question of how simulationists classify cryptomnesia boils down to the question of whether the episodic construction system aims to represent an event from the subject's personal past. There are two options available to the simulationist here, because the notion of 'aim' can be understood in a personal or subpersonal sense (Michaelian, 2016a). Although Michaelian doesn't actually tell us much about what kind of considerations—i.e., personal or subpersonal—should play a role in fixing the aim of the episodic construction systems in specific situations, both accounts can be coherently endorsed. As we show, however, there are problems for both interpretations of cryptomnesia.

Focusing on a personal-level understanding of 'aim', cryptomnesic states are classified as occurrences of imagining, for in those cases the subject aims to represent a possible or fictional event or object, as opposed to an actual event or object belonging to the personal past. The personal-level way of understanding cryptomnesia might seem to account for all the features of cryptomnesia: PHENOMENOLOGY and CREATIVITY seem to be explained by the subject aiming to imagine. And RETENTION and FAILURE seem to be explained by the fact that the simulation theorist can appeal to stored information in the form of memory traces and source monitoring errors. On the face of it, the personal-level simulationist interpretation appears to provide an adequate explanation of cryptomnesia. When we look deeper, however, things are not quite so clear. A first point is that while such an interpretation of the aims of the system in cryptomnesia is logically possible, it doesn't seem to align well with claims made by simulationists elsewhere, which tend to emphasize a subpersonal, or mechanisms perspective. A related, but more important, point is that despite appearing to

---

[14] See, e.g., Michaelian (2016b, 2020), where some memory errors, such as misremembering and relearning, are taken to be the result of a reliably functioning episodic construction system.

adequately account for cryptomnesia, this appearance is deceptive. While the simulation theory has all the pieces necessary to provide a full explanation, the theory, as it stands, cannot put them together in the correct way to account for the phenomenon. Let us explain.

The simulation theory emphasises remembering's backward-looking rather than forward-looking causal role (Langland-Hassan, 2022a, b). Importantly, for the simulationist, characterizing the backward-looking causal role appeals to notions such as reliability rather than causation or content: the source of the information used in the construction of memory is not important for individuating the state or process as remembering. The simulation theory emphasises the synchronic process at the time of retrieval in generating content, not where that content originates (McCarroll, 2020). In this sense, the fact that cryptomnesia may use information stored in a memory trace doesn't *by itself* make the state or process mnemonic. A fully imaginative state or process, one which does not involve an error like cryptomnesia, can still, on the simulation theory, draw on information stored in traces. That, indeed, is the way in which the system generates all manner of representations, from remembering to all other forms of imagining: 'the episodic construction system – is designed to produce representations of past and future events on the basis of raw materials – stored information – deriving from the subject's experiences' (Michaelian, 2022a, 2023, p. 2).

On this way of thinking about how the system uses content, the cryptomnesic subject is simply imagining. Because the type of content involved in generating a simulation is not important for the individuation of the state, then it is the aim of the subject/system (and the reliability of the process) that fixes the kind of state the subject is in. Cryptomnesia, then, would count as simply imagining, rather than counting as a state that is somehow both imaginative and mnemonic. To draw this out a bit more, consider the case in which the subject tries to imagine and draws on memory traces to do so, but knows that the source of information is from a past event. Even in this case, where there is no metacognitive error, the subject counts as simply imagining. The content of the simulation derives from stored memory traces, but it is the aim of the subject that fixes the state in question, not the *source* of the content. One way of putting this point is that with its focus on the synchronic nature of constructing simulations, the simulation theory has not offered us an account of *unsuccessful imagining*. In our view (see Sect. 5), in contrast, content is one important factor in determining the type of state the subject is entertaining.[15]

It is a system- or mechanism-perspective that provides the framework for the second simulationist explanation of cryptomnesia. On this second option, cryptomnesic states are classified as occurrences of episodic remembering, for despite the subject herself intending to represent a possible event or object, unbeknownst to her, the episodic construction system sets for itself the aim of representing an event from the subject's personal past. We outline some possible worries about this way of understanding cryptomnesia below, but it's worth mentioning here that this approach seems to be most in line with the simulationist project. One distinguishing feature of the simulationist approach, when compared to the causalist approach, is that it acknowledges that there

---

[15] Of course, as we acknowledge, the simulation theorist could in principle appeal to content in determining the nature of a mental state; whether this would be an attractive move for them, or whether it would threaten their synchronic understanding of individuating states like remembering and imagining remains to be seen.

is something erroneous in cryptomnesia. This becomes clear in Michaelian's (2016a) brief discussion of cryptomnesia. According to him, there are two possible explanations of what goes wrong in cryptomnesia. The first is to conceive of cryptomnesia as a *source-monitoring* error. Source monitoring (Johnson et al., 1993) can be understood as 'a metacognitive monitoring process in which the subject attempts to determine the origins of memories in sources such as experience, imagination, and communication by relying on features of their contents' (Michaelian, 2016a, p. 64). On this source-monitoring account of cryptomnesia, 'the subject retrieves a stored representation but, perhaps because it lacks the level of detail characteristic of representations originating in experience, takes it to have originated in imagination' (Michaelian, 2016a, p. 173).

The second explanation is to conceive of cryptomnesia as a *process-monitoring* error. Process monitoring is similar to source monitoring, but instead of monitoring features of the representation to determine its origin, there is monitoring of features of the process in order to determine which form of imagination one is engaged in, such as remembering or mere imagining. On the process-monitoring account of cryptomnesia, 'the subject might, for example, remember something only with difficulty and therefore take himself to be imagining—that is, generating a new representation—rather than retrieving a previously generated representation' (Michaelian, 2016a, p. 173).

On either metacognitive explanation of cryptomnesia the result is that the subject is actually remembering even though she takes herself to be imagining. The system simulates an event or experience from the personal past, but due to a metacognitive failure—either source monitoring or process monitoring—the agent mistakes this for imagining. On the face of it, this seems to provide a satisfactory account of cryptomnesia. Both accounts acknowledge that cryptomnesia has a 'mnemonic' dimension and an 'imaginative' dimension. The former aspect is explained by the subpersonal aim of the system, which retrieves a retained content, or reliably constructs an experience from the personal past. The second aspect is explained as a metacognitive failure in which the subject mistakenly believes that she is imagining. But even though these two potential simulationist explanations might be richer and more developed than the one offered by the causal theory, they are not fully satisfactory.

To see this, let us consider them in relation to the characterization of cryptomnesia offered in Sect. 2. RETENTION appears to be unproblematic. Both the source-monitoring and the process-monitoring accounts seem to accommodate it. But, consider FAILURE. While the source-monitoring account also seems to accommodate it—i.e., the idea that subjects fail to identify the source of their representation—it is unclear whether the process-monitoring account is equally successful here. Indeed, on the process-monitoring account, there is some sort of error or failure, but that error or failure doesn't have to do with the source of the information, but rather with the subject's inability to identify the relevant process as one of remembering. So, the process-monitoring account appears to mischaracterize what goes wrong in cryptomnesia.

Consider PHENOMENOLOGY next. The process-monitoring account seems to be able to make sense of it, for it implies that subjects take themselves to be imagining. It is unclear, however, whether the source-monitoring explanation can account for it. The suggestion behind the source-monitoring account seems to be that cryptomnesia involves a source-monitoring error because the remembered information is incorrectly

attributed to imagination. But even if that is the case, it doesn't follow from this that the subject will have an experience of imagining as a result. To see this, consider a scenario in which a subject represents an event that was actually experienced. Suppose, however, that there is a source-monitoring error, such that this representation is incorrectly attributed to an event that the subject imagined in the past. In such cases, the resulting experience is not an experience of imagining, but rather an experience of remembering an imagined event. That source-monitoring does not give us an explanation of how subjects solve the process problem is expected given how source-monitoring works. As Michaelian ([2016a](#), p. 171) writes,

> The source-monitoring framework provides an elegant explanation both of how agents distinguish between representations originating in experience and representations originating in other sources, including imagination, and of why they sometimes fail to do so. It does not, however, explain how agents solve the process problem—nor is it designed to do so, despite occasional ambiguity on this point in the literature.

It might be replied here that perhaps the source error involved in cryptomnesia is one that mistakenly identifies information previously experienced as having its source in the subject's current state, which would explain why the subject take herself to be imagining. But even if that is the case, we are still owed an explanation of why such an attribution suffices for the subject to solve the process problem in these specific cases, while it typically does not in the majority of cases (e.g., when one remembers a past imagined event). In other words, accounting for imaginative phenomenology in terms of source-monitoring threatens to collapse the very distinction between the process problem and the source problem, a distinction that is central for simulationists.

Consider, finally, CREATIVITY. This is where we believe both source-monitoring and process-monitoring explanations are equally problematic. On both explanations, which are based on the subpersonal interpretation of the simulation theory, the cryptomnesic subject is remembering simpliciter. What these explanations imply is that subjects merely *mistake* what they are doing for an imaginative act. But this doesn't imply that the state or process is imaginative. Importantly, this makes it difficult to see how it could be creative in the way specified by CREATIVITY. As we noted before, cryptomnesia often involves modifying mnemonic content, where the retained content is not simply repeated or reconstructed, but is rather embellished and modified. For example, songs that are 'copied' in cryptomnesia are not merely plagiarized but are imaginatively re-invented.

Indeed, this way of thinking about cryptomnesia may be related to how the term was first originally understood. The word appears to be first used by the psychiatrist Théodore Flournoy, in reference to the case of a psychic medium, to suggest a high incidence of 'latent memories on the part of the medium that come out, sometimes greatly disfigured by a subliminal work of imagination or reasoning, as so often happens in our ordinary dreams' ([1901](#)/[1994](#), p. 8). It is not the case that the cryptomnesic subject is simply remembering. Rather there is something both imaginative and mnemonic about the process. Thus, while it may appear at first glance that the source/process monitoring explanation can accommodate these two dimensions, it ultimately fails

to make sense of all the features that make cryptomnesia both an imaginative and a mnemonic phenomenon.[16]

Despite our critical assessment of the simulation theory, we think that it does get something right about cryptomnesia. We think that the simulationist is right in pointing out that there is some metacognitive error involved cryptomnesia. After all, as FAILURE shows, subjects in such states are not aware that the content they entertain is based on some experience they had in the past. Hence, cryptomnesia does indeed seem to involve a source monitoring error. Nonetheless, we need more than this. We think that a better explanation of cryptomnesia will account for all four features that make it both an imaginative and a mnemonic phenomenon, and account for both its backward- and forward-looking causal roles.

5 A three-factor account of cryptomnesia

There is a common way of characterizing mental states in philosophy of mind, which consists in distinguishing between their *contents* and their *attitudes* (Searle, 1983). The content is what the mental state represents, whereas the attitude is the stance taken toward what is represented. This distinction between attitude and content helps explain why we can have *different* mental states, which play different roles in our cognitive economies, that have the *same* content. Remembering and imagining are intentional states—they are states that represent or are about some object or state of affairs. While it is important to think of the relation between them in terms of content, and whether their content is mnemonic, in the sense, for example, that it is retained from previous experience or involves a causal connection to a past event, it is also essential to adequately characterize their respective attitudes (Robins, 2020a; Sant'Anna, 2021).

One helpful way to think about the notion of an attitude in mental states is suggested by Langland-Hassan (2015): '[a] rough-and-ready way to conceive of attitude […] is simply that aspect of a mental state's typical functional role that cannot be accounted for by its content' (p. 667). Thus, for instance, if one thinks that part of what it means to remember is to entertain a content as past, but nonetheless thinks that the content of remembering does not include any temporal information (e.g., De Brigard & Gessell, 2016; Mahr & Csibra, 2018)—perhaps the content is simply a mental image of an

---

[16] The simulationist may be able to offer such an account by, for example, suggesting that cryptomnesia involves a form of memory error (e.g., misremembering) or another form of unsuccessful remembering. This might account for the changes in content that we see in cryptomnesia. Importantly, however, this would still not be a complete explanation of CREATIVITY. Recall that CREATIVITY is still present in cases in which there is an almost verbatim or perfect copy of a previous work that is involved in cryptomnesia, and this means that what we need is not just an explanation of the (potential) change in content but the *experience* that one has created content anew whether or not this is in fact the case. Metacognition may go some way to explaining this aspect of the phenomenology but it cannot provide a full explanation (see Sect. 3). Indeed, memory errors typically come with a phenomenology of *remembering*, not imagining. When one is confabulating or misremembering one still takes oneself to be remembering. Perhaps the simulation theory has the resources to provide a full account of the creativity of cryptomnesia but it hasn't yet done so. And, importantly, the idea that cryptomnesia is a memory error with changed content, which is usually associated with the phenomenology of remembering, while simultaneously involving the phenomenology of imagining, appears to involve a tension.

event—one can account for the "past orientation" of remembering by claiming that its attitude is such that, when one entertains the relevant contents under that attitude, one takes those contents to represent events in the past (Robins, 2020a).[17] Similarly, many of our imaginings are thought to involve a particular attitude, which makes them distinct from remembering. Those imaginings are often characterized in terms of taking an attitude towards a content in an 'epistemically safe' way, such that one is 'engaging in rich, elaborated cognition about the possible, fantastical, pretended, and so on, that is *epistemically compatible* with things not really being the way they are being thought about, and with one's not believing things to be that way' (Langland-Hassan, 2020, p. 5). To imagine, in this sense, 'is to represent without aiming at things as they actually, presently, and subjectively are' (Liao & Gendler, 2019).

How can this distinction between content and attitude be applied to the case of cryptomnesia? Cryptomnesia, we suggest, can be characterised as involving an imaginative attitude towards a retained content, where retained content can be understood as content that is stored in a memory trace—a trace that retains information that one previously encountered, which links back (either directly or indirectly) to one's previous experience. The trace may not involve an appropriate causal connection, but it still provides a link to one's previous experience. Cryptomnesia, then, involves adopting an attitude that does *not* treat the content as having actually occurred in the past, even though the content is one that has in fact occurred in the past. This way of thinking about cryptomnesia provides a full explanation of its 'imaginative' dimension. In particular, if we take cryptomnesia to involve an attitude of imagining, we can explain both PHENOMENOLOGY and CREATIVITY. The reason why subjects have a subjective experience of imagining in cryptomnesia is that those states involve an attitude of imagining. Similarly, the reason why they engage with those contents, making alterations to them as they see fit, is that those contents are entertained as being possible but not actual. In other words, cryptomnesic states don't have a mind-to-world direction of fit, so their contents are not constrained by external states of affairs in the same way that, e.g., the contents of memory are.

It is important to note that not all imaginings will involve what we're calling an 'imaginative attitude'. For instance, some imaginings—call them *actuality-oriented imaginings* (Munro, 2021)—will involve entertaining a content as being actual, such as when one imagines how the Waterloo Battle went or the layout of one's favourite restaurant. Similarly, some other imaginings—call them *imagistic imaginings* (Langland-Hassan, 2020; Van Leeuwen, 2013, 2014)—will involve entertaining a content under no specific attitude, such as when one merely entertains a mental image of an object in mind. As long as those are taken to be genuine occurrences of imagining, it does not come as a surprise that there are different ways in which we can experience a mental state as being an imagining, and hence that there are potentially different ways in which we could try to account for the imaginative dimension of cryptomnesia. So, one natural question is why focus on cases of imagining that involve an

---

[17] There are, of course, concerns with this characterization—e.g., there are cases where we entertain contents as being past that are not occurrences of remembering—but it is not our goal to defend it here. This is just meant to be an illustration of what it means to say that remembering is an attitude. So, while there is room to dispute how we should characterize the attitude of remembering, speaking of it as involving an attitude is not particularly mysterious or unmotivated.

attitude of entertaining a content as being possible or fictional—i.e., call these *attitudinal imaginings* (Langland-Hassan, 2020; Van Leeuwen, 2013, 2014)—when trying to account for cryptomnesia? Wouldn't appealing to the other types of imagining do the trick just as well?

Our focus on cases of attitudinal imaginings is justified by the fact that it is the only type of imagining that can explain both PHENOMENOLOGY and CREATIVITY. While appealing to actuality-oriented and imagistic imaginings might give us an explanation of PHENOMENOLOGY—i.e., why we experience cryptomnesia as an imaginative state—they fail to account for CREATIVITY—i.e., why we experience the contents of cryptomnesia as creative products, such that they can be changed or altered by us as we see fit. Actuality-oriented imaginings fail to explain CREATIVITY because the contents of those imaginings are entertained as actual, meaning that those contents are constrained by the world. Thus, insofar as we continue to entertain those contents as actual, we cannot change them as we see fit. Similarly, imagistic imaginings fail to explain CREATIVITY because entertaining a content as being novel, which is required by CREATIVITY, requires entertaining it in a certain way. And given that entertaining a content as being possible or fictional is at least necessary for entertaining it as being novel, attitudinal imaginings do a better job of explaining CREATIVITY than the other two forms of imagining.[18]

Now, despite successfully capturing the 'imaginative' dimension of cryptomnesia, the appeal to attitudes and content can only provide a partial explanation of its 'mnemonic' dimension. The fact that cryptomnesia involves retained content can give us an explanation of RETENTION, but it does not help us to make sense of FAILURE. This is where we believe that metacognition—in particular source-monitoring—is fundamental to fully account for cryptomnesia. According to our view, states of cryptomnesia are states that involve an imaginative attitude towards a retained content where the source of the latter is incorrectly attributed (by means of a source-monitoring process) to our current state of imagining. The appeal to source-monitoring processes thus makes our account a *three-factor* one, that is, one in which the attitude of the state involved, the nature of the content, and source-monitoring evaluations are all required to determine whether a subject is remembering, imagining, or in a state of cryptomnesia.

---

[18] One concern with this characterization of the imaginative dimension of cryptomnesia is that it fails to fully capture the phenomenology of those states. In particular, it might be argued that the contents of cryptomnesia are not entertained as fictional or possible at all times in cryptomnesia. That is, when a songwriter composes a song, there appears to be a moment in the creative process when the song ceases to be entertained as fictional or possible and is entertained as something 'real'—i.e., it becomes one's *own* composition. In response, it's unclear to us whether that is the case. While it's true that there is a moment in the creative process that a song ceases to be entertained as just a combination of words or sentences and is entertained as a *thing*—i.e., a song—it doesn't seem to follow from this that the song is entertained as 'real', at least not in the sense that it is entertained as being a part of the mind-independent world. This becomes clear when we consider the fact that even after a collection of sentences is entertained as a song during the creative process, the contents of the song are still under the subject's *control*—a feature which, some philosophers have argued, is characteristic of the phenomenology of imagining (Arcangeli, 2018, p. 76; see Kind, 2020 for discussion). It is difficult to see, however, how that could be the case if the song were not entertained as being possible or not real. Thus, if by 'real' we mean that the song becomes a part of the mind-independent world, then it is not true that the objects of cryptomnesia are entertained as real.

This way of thinking about cryptomnesia, and of remembering and imagining more generally, points to a couple of objections that might occur to some readers. First, it might be argued that recent theorizing on the nature of remembering conceives of remembering and imagining as processes, while our focus when we speak of attitudes seems to be on the mental states, or the outputs of those processes. Thus, it could be objected that we are changing the subject of explanation. We don't think this is right. Recent theorizing on remembering in philosophy has largely ignored the mental states that subjects enjoy when they are said to be remembering. This is illustrated, for instance, in the acceptance by both the causal theory and the simulation theory, that the painter in Martin and Deutscher's (1966) example is remembering. By focusing on attitudes, we deny that this is all there is to the *process* of remembering. On our view, the mental state that a subject enjoys on a given situation is fundamental in determining whether they are remembering, imagining, or in a state of cryptomnesia. But this is not to change the subject of the conversation, but rather to directly challenge the more established way of conceiving of remembering, which ignores the subjective dimension and focuses on objective factors to account for remembering understood as a process.

Second, it might be argued that one can remember a content—e.g., a famous passage from a play or book—but experience the relevant state as an imagining because one *believes* that one is imagining. And this is problematic because it allegedly shows that entertaining a content under the attitude of imagining is not necessary for experiencing a mental state as one of imagining.[19] There are two things we want to say in response. The first is that the claim that one can count as *remembering* while *believing* that one is imagining is controversial one. More importantly, it is a claim that we explicitly call into question in our discussion. Our main criticism of the causal and simulation theories is that they fail to incorporate subjective conditions in their analyses of remembering. In particular, we argue that this is a problem because both causal and simulation theories classify as remembering cases in which subjects fail to entertain the contents of their mental states as originating in the past—e.g., the painter case, cryptomnesia. So, we deny that there can be cases of remembering where one takes oneself to be imagining because those fail to satisfy the relevant subjective condition that we think is (partly) defining of remembering.[20] And this is why cryptomnesia is a puzzling phenomenon. If cases of the type just described were to be correctly described as cases of remembering, then cryptomnesia would be *just* remembering. However, as we've argued, cryptomnesia is not just remembering. It has an imaginative dimension that needs to be accounted for.

The second thing we want to say in response is that we also find the claim that "believing that one is φ-ing" is sufficient for one to have an "experience of φ-ing" to be a controversial one. To use the same example introduced above, suppose that after entertaining a passage from a book in one's mind and mistakenly believing that to be a product of one's imagination, one is corrected by a friend who points out that

---

[19] We are grateful to an anonymous referee for raising this objection.

[20] Note, importantly, that our claim here is not that there can't be cases in which one takes oneself to be imagining but is merely *retrieving* information acquired in the past (see the discussion of 'imaginative projections' below), but rather that retrieval of information acquired in the past is not the same as remembering.

the passage is, in fact, from an existing piece of work. Plausibly enough, the subject forms the belief that s/he is remembering. The fact that the subject forms such a belief does not, however, imply that s/he experiences his or her mental state as a memory. According to our three-factor view, for that to be the case, the subject must entertain that content under the attitude of remembering (see below), and having a belief along the lines just described is not sufficient for that to be the case. We think the same is true of imagining. The fact that the subject believes that s/he is imagining is not sufficient for him or her to experience it as an imagining. What is required is entertaining the content in the right way, i.e., under the attitude of imagining.[21] Thus, we don't think that this objection succeeds in dismissing the three-factor view.

In fact, the three-factor view has explanatory breadth. In addition to accounting for cryptomnesia, this three-factor way of thinking also helps us to explain a range of mnemonic and imaginative phenomena. While it is not our goal to fully develop the three-factor account here, briefly discussing its implication for theorizing about the nature of remembering and imagining more generally will help situate our account of cryptomnesia in relation to ongoing discussions in the philosophy of memory and the philosophy of imagination.

Consider remembering first. The three-factor account says that a subject *successfully remembers* when they adopt the attitude of remembering towards a retained content and there is no source-monitoring error involved. This, in turn, provides us with an account of cases of unsuccessful remembering, such as *misremembering* and *false remembering*. In cases of misremembering, where the content of one's memory is partially inaccurate, the three-factor account says that a subject adopts an attitude of remembering towards a modified retained content and there is no a source-monitoring error involved. In cases of false remembering, there are two possibilities. On the one hand, a subject might falsely remember an event when they adopt an attitude of remembering towards a retained content and there is a source-monitoring error involved—for instance, when one remembers an event that was actually experienced by oneself, but mistakenly attributes its source to a dreamed experience. On the other hand, a subject might falsely remember an event when they adopt an attitude of remembering towards non-retained content and there is a source-monitoring error involved—for instance, when one seemingly remembers an event, but the event in question did not happen or one did not experience the event.[22]

---

[21] One possible response to this argument is to say that remembering and imagining are alike in this respect because they involve attitudes of different types. In particular, it could be argued that there is more to the attitude of remembering than believing that one is remembering—which would explain why having this belief is not sufficient for one to experience a mental state as one of remembering—but that the attitude of imagining is *just* the attitude of believing that one is imagining. This is indeed a possibility, but it is at best a controversial one. As we discussed above, it goes against standard ways of characterizing the attitude of imagining, which involve entertaining contents as being possible, fictional, or actual. Moreover, it is unclear what the motivation for this view would be. Claiming that the attitude of imagining is merely the attitude of believing that one is imagining says very little, if anything at all, about what it is to imagine or what it means to entertain contents in an imaginative way.

[22] Due to space limitations, we won't discuss cases of *confabulation* here (Hirstein, 2005). We believe, however, that the three-factor account can make sense of them. How it does so will depend on whether we take confabulations to be a fundamentally *clinical* phenomenon. While the standard approach has been to treat them as such (e.g., Hirstein, 2005; Michaelian, 2016b, 2020), Robins (2020b) has recently argued that there is an important distinction to be made between 'mnemonic' and 'clinical' confabulations. If we

Consider *imagining* now. The three-factor account allows us to distinguish between three types of imagining. A first type is what we call *pure creative imaginings*: when a subject adopts an attitude of imagining towards a non-retained content and there is no source-error involved. For instance, when a musician composes a completely new song.[23] A second type is what we call *mnemonic imaginings*: when a subject adopts an attitude of imagining towards a modified retained content and there is no source-error involved. For instance, when one imagines the layout of one's apartment and uses that as a starting point for imagining the layout of a new apartment. Finally, a third type is what we call *imaginative projections*: when a subject adopts an attitude of imagining towards a retained content and there is no source-error involved. For instance, when one is asked to imagine the next time one will go to the beach, one does so by remembering the last time one went to the beach and projecting that representation into the future.

To briefly illustrate the advantages of the three-factor account, consider how it helps us to explain cases of imaginative projection. Imaginative projections, as we noted above, are cases in which we project a retained content into the future, such as when one is asked to imagine a future trip to the beach and does so by simply recalling a previous experience of going to the beach and entertaining it as a future event. Unlike cryptomnesia, these cases don't involve a source-error since subjects are aware that they are imagining on the basis of a previous experience. However, like cryptomnesia, there is an 'imaginative' dimension and a 'mnemonic' dimension to this phenomenon. Subjectively speaking, the subject takes herself to be imagining the event in question. At the level of the operating mechanisms, though, those imaginings draw on retained contents. The three-factor account is able to make sense of the 'hybrid' nature of these states.

However, both the causal theory and the simulation theory face trouble when trying to account for imaginative projections. Since imaginative projections involve retained content, which in the context of the causal theory means that they stand in appropriate causal connections to past experiences, they will be classified as cases of remembering *tout court*. But this is problematic, for it doesn't adequately account for what the subject takes herself to be doing, nor the state's forward-looking causal role.[24]

The simulation theory faces similar problems. Assuming that the reliability condition is satisfied in imaginative projections, there seems to be two options available to simulationists. The first is to say that the first 'aim' of the system was to represent the past, so the subject is remembering. But then we have a case of remembering the future, or future-oriented remembering because the subject is using that representation to think about the future. Again, just like the causal theory's classification, this seems puzzling. The second option is to say that it is the second 'aim' of the system that is

Footnote 22 continued

understand mnemonic confabulations as being fundamentally inaccurate representations of past events under a remembering attitude in which no malfunctioning is involved—i.e., as a non-clinical phenomenon—then they will count as cases of false remembering where (false) content is retained and where there is a source error. If, however, we understand mnemonic confabulations as involving malfunctioning—i.e., as a clinical phenomenon—then they will count as cases of false remembering where content is *not* retained and where there is a source error.

[23] This is compatible with cases in which the subject draws on retained semantic information.

[24] See Sect. 6 for further discussion of imaginative projections, some of which may involve planning for the future.

important (the future event). In this case the subject is aiming to imagine the future and hence there is no incoherence. But then, we suggest, the simulation theory fails to account for how this form of imaginative projection is *both* imagining and mnemonic. Again, the simulation theory seems to offer an either/or account of this type of phenomenon, which doesn't fully explain the imaginative and mnemonic dimensions of these processes.

6 Cryptomnesia and epistemic relevance

Philosophical discussions of cryptomnesia are rare. Nonetheless, the painter case (Martin & Deutscher, 1966) is a broadly similar phenomenon. As we saw previously, Martin and Deutscher think that the painter is remembering. We rejected this interpretation because it fails to acknowledge the personal level aspect of the phenomenon. In fact, there are other accounts that also reject the causal theory's interpretation of the phenomenon (Debus, 2010; Fernández, 2019).[25]

According to Debus (2010), the painter is not remembering even though the causal theory's three conditions on successful remembering are satisfied. The three conditions might be necessary, but they are not sufficient for genuine remembering. What is missing in the painter case that rules out genuine remembering? The answer, according to Debus, is *epistemic relevance* for the subject. Epistemic relevance refers to the idea that experiential (episodic) memories usually 'play a role in our forming judgements about the past' (Debus, 2010, p. 20). Epistemic relevance means that the subject is disposed to take a present experience (an occurrent memory) into account when making judgements about the past. Present experiences such as memories have epistemic relevance for the subject when they reconfirm an existing belief, lead the subject to reconsider an existing belief, or provide the subject with new knowledge about the past. How does our view relate to this causal-epistemic position?

While we think that the causal-epistemic position gets it right that the painter, and other cryptomnesic subjects, are *not* remembering, we think that such accounts don't provide the full picture. The causal-epistemic view provides an account of cryptomnesia in which it is characterised negatively: such accounts tell us what type of state cryptomnesia *is not*; that is, they tell us that cryptomnesia is *not* remembering. Such accounts, however, do not offer a positive account and specify what cryptomnesia *is*. The causal-epistemic view doesn't explain the imaginative component of cryptomnesia and how the phenomenon may be simultaneously imaginative and mnemonic. Indeed, it doesn't explain how the mnemonic content may be imaginatively altered or embellished rather than simply reproduced. In contrast, our view neatly explains the imaginative aspect of cryptomnesia.

Indeed, arguably our view explains why episodic or experiential memories typically do play a role in forming judgments about the past etc. That is, our view, with

---

[25] In what follows we focus on Debus's account of 'epistemic relevance', which is a condition that rules out remembering in the type of cases exemplified by the painter. Robins (2020b, fn. 3) also denies that the painter case involves remembering. The reason is that, for Robins, successful (and unsuccessful) remembering involve an attitude of *seeming to remember*, and this is absent in cryptomnesic cases (see also Robins, 2020a).

its emphasis on the attitude that one can take towards various types of content, helps explain why memories are epistemically relevant to the subject: one adopts an attitude of remembering (or seeming to remember if we want to be open to non-factive instances of remembering) when one takes the content of one's occurrent mental state to be targeting an experience from one's personal past (Robins, 2020a, b). It is the attitude towards the content that arguably links the experience with beliefs and enables it to play a role in making judgements about the past. After all, one can adopt a different attitude (e.g., counterfactual thinking) towards the same content and this will result in a different set of beliefs or judgements about the event. In fact, our account elucidates why cryptomnesia does not result in beliefs about the past. On our view, in cryptomnesia the subject adopts an imaginative attitude towards mnemonic content, where this imaginative attitude 'is *epistemically compatible* with things not really being the way they are being thought about, and with one's not believing things to be that way' (Langland-Hassan, 2020, p. 5). This imaginative attitude helps explain why cryptomnesic states are not epistemically relevant to the subject, even though they involve content that (causally) derives from one's past experience.

Perhaps one might object here that there is a tension in the way in which we are employing the notion of an imaginative attitude in explanations of epistemic irrelevance. We suggest that because cryptomnesia involves an imaginative attitude, this explains the way in which the content is epistemically irrelevant for the subject. We also suggest that this same imaginative attitude may be involved in cases of imaginative projections, in which we project mnemonic content into a future scenario. The tension arises because imaginative projections might form the basis for judgments about the future, and hence the imaginative attitude is in play in a context where it seems to be epistemically relevant. In other words, we may be equivocating on the sense in which the imaginative attitude renders a state imaginative.[26]

This is an interesting objection, but we think that there might be an important sense in which the tension dissolves when we consider that there might be varieties of imaginative projections, which involve different imaginative attitudes (Langland-Hassan, 2015; see also Sect. 5). The type of case that is supposed to be problematic for our account is the case in which the subject uses imagination to make judgements about the future, in order to predict or plan for it. This is indeed a case in which the imagination is used in an epistemically relevant way. Arguably, however, this type of imaginative project involves an attitude that is more akin to a judgement than the imaginative attitude. But not all imaginative projections need be like this. One could simply imagine some future event in the context of an idle daydream or fantasy. In this context, there is no judgement about the future that one is trying to form. Arguably, the difference between the two forms of imaginative projection—epistemically relevant and epistemically irrelevant—is the type of attitude that one is adopting. In the epistemically relevant case, the attitude is something like an imaginative judgement (Langland-Hassan, 2022, 2023; cf. McCarroll, 2023), whereas in the epistemically irrelevant case the attitude is an imaginative attitude as we have been discussing it in this paper. An epistemically irrelevant imaginative projection involves the same kind of imaginative attitude as is involved in cryptomnesia.

---

[26] We thank an anonymous reviewer for raising this interesting objection and pressing us on this point.

7 Conclusion

Cryptomnesia is a complex phenomenon, which is somehow both simultaneously imaginative and mnemonic. The most prominent theories of episodic memory, the causal theory and the simulation theory, cannot account for all the features of this phenomenon. Instead, we have argued that the best way to fully capture the mnemonic and imaginative aspects of cryptomnesia is to think of it in terms of an attitude and content pairing combined with a source monitoring error. Our three-factor account provides a full explanation of cryptomnesia, accounting for all of the four features that it exhibits. In addition, our three-factor theory sheds light on a range of other mnemonic and imaginative phenomena.

Declarations

Conflict of interest The authors have no conflict of interest to declare.

References

Andonovski, N. (2020). SINGULARISM about episodic memory. *Review of Philosophy and Psychology, 11*(2), 335–365.

Arcangeli, M. (2018). *Supposition and the imaginative realm: A philosophical inquiry*. Routledge.

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

Brown, A. S., & Murphy, D. R. (1989). Cryptomnesia: Delineating inadvertent plagiarism. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(3), 432–442.

Craver, C. F. (2020). Remembering: Epistemic and empirical. *Review of Philosophy and Psychology, 11*(2), 261–281.

De Brigard, F. (2014). The nature of memory traces. *Philosophy Compass, 9*(6), 402–414.

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.

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

Dokic, J. (2014). Feeling the past: A two-tiered account of episodic memory. *Review of Philosophy and Psychology, 5*(3), 413–426.

Dokic, J. (2021). Episodic remembering and affective metacognition. *Acta Scientiarum: Human and Social Sciences, 43*, e61022.

Fernández, J. (2019). *Memory: A self-referential account*. Oxford University Press.

Flournoy, T. (1901/1994). *From India to the Planet Mars: A case of multiple personality with imaginary*. Princeton University Press.

Gerrans, P., & Kennett, J. (2017). Mental time travel, dynamic evaluation, and moral agency. *Mind, 126*, 259–268.

Hirstein, W. (2005). *Brain fiction: Self-deception and the riddle of confabulation*. MIT Press.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114*(1), 3–28.

Kind, A. (2016). Introduction: Exploring imagination. In A. Kind (Ed.), *The Routledge handbook of philosophy of imagination*. Routledge.

Kind, A. (2020). Imaginative experience. In U. Kriegel (Ed.), *The Oxford handbook of the philosophy of consciousness*. Oxford University Press.

Langland-Hassan, P. (2015). Imaginative attitudes. *Philosophy and Phenomenological Research, 90*(3), 664–686.

Langland-Hassan, P. (2020). *Explaining imagination*. Oxford University Press.

Langland-Hassan, P. (2021). What sort of imagining might remembering be? *Journal of the American Philosophical Association, 7*(2), 231–251.

Langland-Hassan, P. (2022). Propping up the causal theory. *Synthese, 200*, 95.

Langland-Hassan, P. (2023). Remembering and imagining: The attitudinal continuity. In A. Berninger & Í. Vendrell Ferran (Eds.), *Philosophical perspectives on memory and imagination*. Routledge.

Liao, S., & Gendler, T. (2019). Imagination. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/archives/win2019/entries/imagination/

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

Marsh, R. L., Landau, J. D., & Hicks, J. L. (1997). Contributions of inadequate source monitoring to unconscious plagiarism during idea generation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(4), 886–897.

Martin, C. B., & Deutscher, M. (1966). Remembering. *Philosophical Review, 75*(2), 161–196.

McCarroll, C. J. (2020). Remembering the personal past: Beyond the boundaries of imagination. *Frontiers in Psychology, 11*, 2652.

McCarroll, C. J. (2023). Memory and imagination, minds and worlds. In A. Berninger & Í. Vendrell Ferran (Eds.), *Philosophical perspectives on memory and imagination* (pp. 35–53). Routledge.

Michaelian, K. (2016a). *Mental time travel: Episodic memory and our knowledge of the personal past*. MIT Press.

Michaelian, K. (2016b). Confabulating, misremembering, relearning: The simulation theory of memory and unsuccessful remembering. *Frontiers in Psychology, 7*, 1857.

Michaelian, K. (2020). Confabulating as unreliable imagining: In defence of the simulationist account of unsuccessful remembering. *Topoi, 39*(1), 133–148.

Michaelian, K. (2022). Radicalizing simulationism: Remembering as imagining the (nonpersonal) past. *Philosophical Psychology*. https://doi.org/10.1080/09515089.2022.2082934

Michaelian, K. (2023). If Remembering is Imagining, then what is Forgetting? In A. Berninger & Í. Vendrell Ferran (Eds.), *Philosophical perspectives on memory and imagination* (pp. 76–93). Routledge.

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.

Munro, D. (2021). Remembering the past and imagining the actual. *Review of Philosophy and Psychology, 12*(2), 175–197.

Oppel, F. N. (2005). *Nietzsche on gender: Beyond man and woman*. University of Virginia Press.

Perrin, D., Michaelian, K., & Sant'Anna, A. (2020). The phenomenology of remembering is an epistemic feeling. *Frontiers in Psychology, 11*, 1531.

Robins, S. K. (2016). Representing the past: Memory traces and the causal theory of memory. *Philosophical Studies, 173*(11), 2993–3013.

Robins, S. K. (2020a). Defending discontinuism, naturally. *Review of Philosophy and Psychology, 11*, 469–486.

Robins, S. K. (2020b). Mnemonic confabulation. *Topoi, 39*(1), 121–132.

Sant'Anna, A. (2021). Attitudes and the (dis)continuity between memory and imagination. *Estudios de Filosofía, 64*, 73–93.

Savage, M. (2016). Dolly Parton defends Ed Sheeran, Led Zeppelin and others accused of theft. https://www.bbc.com/news/entertainment-arts-37118347

Schechtman, M. (2001). Empathic access: The missing ingredient in personal identity. *Philosophical Explorations, 4*(2), 95–111.

Searle, J. R. (1983). *Intentionality: An essay in the philosophy of mind*. Cambridge University Press.

Skinner, B. F. (1983). Intellectual self-management in old age. *American Psychologist, 38*(3), 239–244.

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

Van Leeuwen, N. (2013). The meanings of "imagine" Part I: Constructive imagination. *Philosophy Compass, 8*(3), 220–230.

Van Leeuwen, N. (2014). The meanings of "imagine" Part II: Attitude and action. *Philosophy Compass, 9*(11), 791–802.

Werning, M. (2020). Predicting the past from minimal traces: Episodic memory and its distinction from imagination and preservation. *Review of Philosophy and Psychology, 11*(2), 301–333.

Womack, K. (2016). *The Beatles encyclopedia: Everything Fab Four*. Greenwood.

Yin, S. (2016, August 27). The accidental plagiarist in all of us. *The New York Times*. https://www.nytimes.com/2016/08/27/science/cryptomnesia-demi-lovato-plagiarism.html