Autonoesis and episodicity: Perspectives from philosophy of memory

André Sant’Anna¹,² | Kourken Michaelian²,³ | Nikola Andonovski²

¹CONCEPT, University of Cologne, Cologne, Germany
²Centre for Philosophy of Memory, IPhIG, Université Grenoble Alpes, Saint-Martin-d’Heres, France
³Institut Universitaire de France, France

Correspondence
Kourken Michaelian, Centre for Philosophy of Memory, IPhIG, Université Grenoble Alpes, Saint-Martin-d’Heres, France. Email: kourken.michaelian@univ-grenoble-alpes.fr

Funding information
André Sant’Anna is funded by an Alexander von Humboldt Postdoctoral Fellowship; European Union’s Horizon research and innovation program under the Marie Skłodowska-Curie, Grant/Award Number: 101062754

Abstract
The idea that episodic memory is distinguished from semantic memory by the fact that it involves autonoetic consciousness, initially introduced by Tulving, has been influential not only in psychology but also in philosophy, where a variety of approaches to autonoesis and to its relationship to episodicity have been developed. This article provides a critical review of the available philosophical approaches. Distinguishing among representational, metacognitive, and epistemic accounts of autonoesis, it considers these in relation to objective and subjective conceptions of episodicity and assesses them against immediacy and source criteria that any philosophical account of autonoesis should arguably aim to satisfy.

This article is categorized under:
- Philosophy > Psychological Capacities
- Philosophy > Consciousness
- Psychology > Memory

KEYWORDS
autonoesis, episodic memory, episodicity, phenomenology, philosophy of memory

1 | INTRODUCTION: AUTONOETIC CONSCIOUSNESS AND THE EPISODICITY QUESTION

Within the category of declarative memory, episodic memory is often distinguished from semantic memory by the involvement in episodic remembering of what Tulving (1985a, 1985b, 2002) referred to as autonoetic (or “self-knowing”) consciousness.¹ Unlike semantic and procedural remembering, which he took to be characterized, respectively, by noetic (“knowing”) and anoetic (“unknowing”) consciousness, Tulving saw episodic remembering as characteristically involving a sense that the remembered event belongs to one’s personal past, that one experienced the event first-hand, or perhaps that one’s memory of the event originates in one’s experience of it. Autonoesis—sometimes described by Tulving as a sense of “reliving” or “reexperiencing” an event or as a sense of “mentally travelling in time”—can be seen as an answer to what has been called the episodicity question (Michaelian & Sutton, 2017; Perrin & Rousset, 2014): what distinguishes between episodic memory, on the one hand, and semantic and other forms of memory, on the other?
The episodicity question has drawn the attention not only of psychologists but also of philosophers. In attempting to answer the question, philosophers have often followed Tulving in assigning a central role to the form of conscious experience that is characteristic of episodic remembering—that is, to the phenomenology of memory. This similarity notwithstanding, their understanding of the episodicity question has differed from Tulving’s own understanding. Whereas Tulving’s primary concern was to distinguish between episodic memory and semantic memory understood as systems (e.g., Tulving, 1985a: 385–388, 2002: 5–6), philosophers have been concerned primarily to distinguish between episodic memory and semantic memory understood as conscious or experiential states. Due to this focus on states rather than systems, the accounts of autonoesis that are available in the philosophical literature differ markedly from those on offer in the psychological literature.

This article critically reviews the available philosophical accounts of autonoesis. Section 2 begins by providing some necessary background. Each of the following three sections examines a particular family of accounts of autonoesis. Section 3 looks at representational accounts, which explain autonoesis in terms of the nature of episodic memory representations. Section 4 considers metacognitive accounts, which explain it in terms of metacognitive states involved in episodic remembering. Section 5 turns to epistemic accounts, which explain it by treating episodic memory as a retained epistemic state. Section 6 brings the article to a conclusion by identifying potential avenues for future research on autonoesis.

2 | BACKGROUND

This section clears up a number of terminological ambiguities and sets out criteria that any philosophical account of autonoesis should arguably aim to satisfy.

2.1 | Autonoesis

The terminology employed in philosophical attempts to answer the episodicity question by appealing to the phenomenology of remembering is unsettled. In an older discussion that remains influential today, Russell (1921/1995) suggested that remembering involves both a “feeling of familiarity” and a “feeling of pastness”. In addition to these feelings and to those discussed below—Fernández (2017, 2019), for instance, invokes a “feeling of past experience” apparently meant to be distinct from the feeling of pastness—recent authors have described a number of other feelings, with Klein and Nichols (2012), for example, referring to a “feeling of ownership” for memory (see also Fernández, 2019; Gentry, 2023; Nichols, 2017). It is not clear, however, whether these are merely different terms that are used to refer to the same feature of the phenomenology of episodic memory or whether they refer to different but potentially related features of that phenomenology.

To avoid confusion in this respect, we will follow Tulving and use the term “autonoesis” in our discussion. We will, in particular, take autonoesis to be the feature of the phenomenology of episodic remembering in which episodic memory is experienced as originating in a past experience or event. This definition is fit for our purposes for three reasons. First, it is compatible with the idea that the phenomenology of episodic memory is not exhausted by autonoesis, thus leaving room for a more general account of that phenomenology that makes reference to the feelings discussed above. Second, the definition is broadly in line with Tulving’s definitions of the term in different works, which emphasize the relationship between episodic remembering and a past experience or a past event (e.g., Tulving, 1985b), while remaining neutral on more controversial aspects of his approach, such as the view that autonoesis necessarily involves consciousness of the self or a sense of “mental time travel” (see Tulving, 2002). And third, the definition is neutral with respect to the various accounts surveyed below. As it will become clear in due course, those accounts offer different explanations not only of what it means for episodic memory to be experienced as originating in a past experience or event but also of what aspects of episodic memory understood as an experiential state make it the case that it is experienced in this way. For instance, some accounts hold that autonoesis is a result of episodic memory representing itself as standing in a specific causal or informational relationship to past experiences, while others maintain that this relationship is best characterized as being justificatory or epistemic in nature. We note that we do not take ourselves to be offering a theoretical definition of autonoesis but merely a pragmatic definition that serves reference-fixing purposes in the context of our project.
2.2 | The relationship between episodicity and autonoesis

One important preliminary question concerns the relationship between episodicity and autonoesis. Broadly speaking, episodicity refers to the feature of episodic memories that makes them episodic. Autonoesis, as we define it, refers to the feature of the phenomenology of episodic remembering in which episodic memory is experienced as originating in a past experience or event. When the notions are understood in this way, the projects of explaining autonoesis and answering the episodicity question are, from a strictly logical point of view, separate and independent.

This does not, however, reflect how the relevant discussions have unfolded in the recent literature. Inspired by Tulving, many attempts to account for autonoesis have also been developed with the goal of accounting for episodicity. These accounts aim not only to show why episodic memory is experienced as originating in a past experience or event but also to show that such an experience is not present in semantic memory. In dialectical terms, then, accounts of autonoesis have also been treated, at least in part, as accounts of episodicity. In line with this dialectical trend, we take accounts of autonoesis to be attempts to explain both (1) why episodic memory is experienced as originating in a past experience or event and (2) why such experience is present only in episodic memory.

One important clarification, which may prove helpful in assessing the relevant accounts, concerns the strategy adopted by their proponents in explaining (2). In addressing this question, they have engaged primarily in a positive project, which consists in showing that episodic memory possesses the feature that explains the involvement of autonoesis in it. It is then argued on that basis that, given that semantic memory does not appear to involve autonoesis, it lacks the feature that accounts for the presence of autonoesis in episodic memory. A natural question to ask here, however, is whether we have reasons other than phenomenological data to believe that semantic memory does, as a matter of fact, lack the feature in question. While pursuing this negative project seems like a natural next step in attempts to explain (2), it has only been addressed indirectly in the existing literature, which makes it difficult to identify the relevant arguments and assess their implications. For this reason, our review will focus primarily on the positive project, but we take this clarificatory note as an opportunity to highlight an important dimension in which theorizing about the nature of autonoesis needs to be further developed.

2.3 | Two episodicity questions

An additional potential source of confusion lies in the fact that there are two related but importantly different episodicity questions. The first concerns the difference between episodic memory and semantic memory from a strictly ontological point of view:

The Objective Question

What, objectively speaking, distinguishes episodic memory from semantic memory?

The goal of an account of autonoesis developed with this question in mind will be to explain the objective distinction between episodic memory and semantic memory in phenomenological terms—to show that it is in virtue of the involvement in a given memory of a certain form of consciousness that that memory qualifies as episodic. Such an account will, in other words, treat autonoesis as an objective “episodicity marker.”

The second question concerns the difference between episodic memory and semantic memory from a broadly epistemic point of view:

The Subjective Question

What, from the subject’s point of view, distinguishes episodic memory from semantic memory?

The goal of an account of autonoesis developed with this question in mind will be to explain the subjective distinction between episodic memory and semantic memory in phenomenological terms—to show that it is the involvement in a given memory of a certain form of consciousness that leads the rememberer to judge that that memory is episodic. Such an account will, in other words, treat autonoesis as a subjective “episodicity marker.”

From a strictly logical point of view, the objective and subjective questions are independent of one another. A given account might, for instance, target the objective question while remaining silent about the subjective question. Similarly, a given account might attempt to answer both questions, but develop those answers in ways that are largely
independent of one another. For instance, one might hold that the presence of a certain conscious experience is what leads subjects to judge that a memory is episodic—a judgment that they need not, of course, formulate in those terms—while holding that the factor in virtue of which the memory in question qualifies as episodic is that it is the result of unconscious processes of a certain kind. This is an important clarification, for despite being logically independent, these questions are rarely distinguished in the relevant discussions, with many of the existing accounts taking it for granted that the answer they offer to one of them entails a parallel answer to the other. We will consider this point in more detail in connection with each account surveyed below.

2.4 Criteria

A number of distinct accounts of autonoesis are available in the philosophical literature, but there has been little engagement among the partisans of those accounts, and there is, in consequence, little clarity with respect to how the accounts might be compared to one another. In order to begin to bring the accounts into contact with each other, the following sections will consider them in relation both to the two episodicity questions described above and to a pair of criteria that any philosophical account of autonoesis arguably ought to satisfy.

The first of these, the immediacy criterion, says that an account ought to do justice to the fact that our awareness of an episodic memory as an episodic memory is a direct or immediate result of our experiencing autonoesis, as opposed to an inferential act or an additional act of awareness occurring subsequent to or in parallel with autonoesis. An account of autonoesis ought to get this feature right insofar as one of its goals is to accurately describe the phenomenology of episodic memory. The second criterion, the source criterion, says that an account ought to provide at least a general indication of the cognitive process that underwrites autonoesis. Satisfying the source criterion is desirable insofar as a theory must be broadly compatible with developments in the empirical sciences in order to be acceptable.

Note that the focus of the article on these criteria is meant to suggest neither that they are the only ones that might be relevant nor that an account must meet both in order to be acceptable; the suggestion is merely that they constitute useful starting points for a comparative assessment of the available accounts of autonoesis. In this sense, it is an open question how proponents of the accounts surveyed below will try to make sense of them. As we will see, different accounts struggle to satisfy at least one of these criteria. One natural way to deal with these struggles is for those accounts to be developed in a way that accommodates the relevant criterion. Another is to reject, or at least to cast doubt on, the necessity of satisfying the relevant criterion. For instance, it might be argued that expecting accounts of autonoesis to satisfy both the immediacy and source criteria is too strong a requirement, one that is implausible in light of the methodological commitments that drive different philosophical projects. We intend to remain neutral on these methodological issues, as they have not been discussed explicitly in the existing literature. Given that our primary goal is to offer a framework against which to compare these accounts, and given that existing accounts have attempted to satisfy at least one of the criteria, we take the assumption that a comprehensive account of autonoesis arguably ought to satisfy both criteria to be a reasonable starting point.

3 REPRESENTATIONAL ACCOUNTS

Representational accounts, as noted above, attempt to explain autonoesis in terms of the nature of episodic memory representations.

3.1 Types of representational account

Representations are standardly understood as involving both attitudes and contents. Two types of representational account are thus available. The first attempts to explain autonoesis in terms of the involvement in episodic memory representations of a distinctive kind of content. The second attempts to explain it in terms of the involvement of a distinctive kind of attitude.
3.1.1 | Content-based accounts

Several content-based accounts have been proposed. One prominent such account has been articulated independently by Dokic (2001) and Fernández (2006, 2017, 2019), who attribute reflexive or self-referential contents to episodic memory. An episodic memory, on Dokic’s and Fernández’ view, represents itself as being caused by the subject’s past experience. It is because the memory has this self-referential content—content that pertains to its own relation to past experience—that the experience of episodic remembering differs from that of semantic remembering. Semantic remembering can make one aware that an event happened in the past and that it was previously experienced by one, but because it does not include self-referential content, it does not make one aware of an event as something that happened in the past or as something that one previously experienced.

A second content-based account has been proposed by Rowlands (2018), who argues that episodic memory involves Fregean modes of presentation (that is, specific ways in which information is presented to the subject in consciousness). An episodic memory, on Rowlands’ view, not only contains information about an event but also presents that information to consciousness in a unique way, namely, as originating in the subject’s past experience. While Rowlands’ primary concern is to explain how, in episodic remembering, one becomes immediately aware of a memory’s origin or source—he contrasts the experience of remembering to the experience of looking at a photograph of a past event, in which one is required to perform an inferential act to become aware of a past experience as its source—his characterization of the content of episodic memory serves equally well to distinguish episodic remembering from semantic remembering: because the contents of semantic memories do not involve the mode of presentation that is involved in the content of episodic memories, they do not make one aware of events as events that one experienced.

A third content-based account has been defended by Hoerl (2001), who holds that episodic memory involves a distinctive kind of causal understanding of the world, a kind of understanding that is bound up with its distinctive phenomenology. This phenomenology is, on Hoerl’s view, to be explained in terms of the spatial contents of episodic memories. Drawing on an analogy to perception, which he takes to involve an awareness of its own “spatial enabling conditions” (2001: 332), Hoerl argues that, in episodic remembering, one is aware of the spatial enabling conditions of a past experience. It is in virtue of this feature that, on Hoerl’s view, episodic remembering differs from semantic remembering at the level of consciousness: when one remembers semantically, one becomes aware of what happened, but one does not become aware of the particular way in which one came to know what happened or of what the causal constraints were on one’s acquisition of that knowledge.

One thing that is worth noting about content-based accounts is that they seem to align well with existing representationalist or intentionalist accounts of consciousness, according to which the phenomenology of mental states is explained by what those states represent (see Lycan, 2019, Pautz, 2020 for overviews). This point is made explicitly by Fernández (2019), who takes this to be an important motivation for his approach. While other content-based theorists have not been as clear on this issue, their accounts also seem to be compatible with representationalist or intentionalist approaches. Given the popularity enjoyed by these approaches in recent philosophy of mind, one might thus view their alignment with content-based accounts as an argument in their favor.

3.1.2 | An attitude-based account

An attitude-based alternative to content-based representational accounts has been developed by Mahr and Csibra (2018). Mahr and Csibra’s account is closely related to their view of the function of episodic memory. They argue that episodic memory functions to regulate communication about the past in two ways: first, by enabling speakers to provide hearers with reasons to accept their reports about the past; second, by enabling hearers to represent reasons to accept or reject speakers’ reports about the past. It is this capacity to represent reasons that Mahr and Csibra take to explain autonoesis. Borrowing the concept of an “epistemic attitude” of remembering—which can be characterized, roughly, as the attitude that one adopts when one takes information to have been obtained first-hand—from Cosmides and Tooby (2000), Mahr and Csibra argue that autonoesis is “an outcome of the capacity to metarepresentationally embed outputs of the scenario construction system [i.e., event representations] under the epistemic attitude of remembering” (2018: 4). In concrete terms, what this means is that, when we remember, we represent the justificatory relationship between beliefs formed on the basis of memory and their experiential sources. In other words, we represent the reasons we have for believing the things we do on the basis of memory (2018: 9). The involvement of this attitude in episodic
remembering is, in turn, what distinguishes episodic from semantic remembering, for, while semantic remembering does allow one to access information about the past, it does not provide one with reasons for believing that information.

There has not been much discussion about the relationship between content- and attitude-based accounts in the literature, but a key factor in choosing between these types of account will be the strategy that one adopts to individuate the content of episodic memory. If content is individuated in terms of our intuitions regarding the accuracy conditions of remembering or introspective reflection on consciousness, which seems to be the strategy employed by Fernández, Dokic, and Rowlands, then we seem to have good reasons for thinking that autonoesis is part of the content of remembering and not of the attitude of remembering. If, however, content is individuated in terms of the outputs of what we might call an episodic construction system (Michaelian, 2016), the function of which is to produce representations of events that are entertained not only in episodic memory, but also in other forms of episodic imagination—a strategy that Mahr and Csibra (2018) appear to endorse—then there is good reason for thinking that autonoesis is, rather, a result of the way or attitude in which that content is entertained. What the advantages and disadvantages of each of these strategies are is a question that future work by partisans of the representationalist approach should consider.

3.2 Assessing representational accounts

Attitude- and content-based representational accounts are situated similarly with respect to the objective and subjective episodicity questions. They fare similarly with respect to the immediacy criterion but appear to fare differently with respect to the source criterion.

3.2.1 Objective and subjective episodicity

Because they explain autonoesis in terms of the representational structure of episodic memory, representational accounts can be plausibly interpreted as attempts to answer primarily the objective question. Since they hold that what it means for an experiential state to be an episodic memory is for it to have a certain representational structure, and since having that representational structure is just what it is for an experiential state to be autonoetic, they seem to be committed to the claim that being an episodic memory just is being an autonoetic experiential state. More specifically, according to content-based accounts, an experiential state counts as an episodic memory when it involves a certain type of content. According to attitude-based accounts, an experiential state counts as an episodic memory when it involves a certain type of attitude.

It is less clear, however, whether they should also be interpreted as attempts to answer the subjective question. Although it would seem like a natural next step for proponents of representational accounts to claim that subjects are able to identify episodic memories as being episodic in virtue of their having conscious experiences of a certain type—either because those memories involve a specific type of content or because they involve a specific type of attitude—this is not a point that has been explicitly discussed in the context of representational accounts. If, however, we assume that representational accounts are indeed attempts to answer the subjective question, then content-based accounts seem to imply that subjects judge that they are remembering because their experiential states involve self-referential phenomenology (Fernández, Dokic, and Rowlands) or because they are aware of the enabling conditions of past experiences (Hoerl). Likewise, attitude-based accounts seem to imply that judgments concerning the episodicity of memory result from subjects becoming aware of the justificatory relationship between perceptual and mnemonic beliefs. Which of these accounts better makes sense of our capacity to make episodicity judgments is an interesting open question.

3.2.2 Phenomenology and source

Both content-based and attitude-based representational accounts appear to be well suited to explain the immediacy of our awareness of an episodic memory as an episodic memory. According to content-based accounts, having an episodic memory is a matter of having a mental state with a specific sort of content. Since all it takes for one to experience autonoesis is, if these accounts are right, for one to have a mental state with the sort of content in question, there is no additional inferential act or act of awareness required for one to become aware of an episodic memory as an episodic
memory. According to attitude-based accounts, having an episodic memory is a matter of embedding a content under a specific sort of attitude. Since the process of embedding a content under an attitude is not something that the subject does—that is, it is not a personal-level process—but is, rather, the result of the operation of subpersonal metarepresentational processes at retrieval, there is, again, no additional inferential act or act of awareness required for one to become aware of an episodic memory as an episodic memory.

While attitude- and content-based accounts thus fare similarly with respect to the immediacy criterion, they may fare differently with respect to the source criterion. Mahr and Csibra, in developing their attitude-based account, are motivated by empirical considerations pertaining to the operation of metarepresentational processes at retrieval. Thus, while the account does not yet provide a full-blown description of the cognitive processes responsible for generating autonoesis, it is well positioned to do so. In contrast, Dokic, Fernandez, Rowlands, and Hoerl, in developing their content-based accounts, are motivated by more purely philosophical concerns and thus have had little to say about the source of autonoesis. Thus, in their existing versions, their accounts seem to be poorly positioned with respect to the goal of accounting for the generation of autonoesis.

Proponents of content-based accounts might respond here by saying that this is not a limitation of their accounts, but only a difference of priority, in the sense that their efforts have so far been focused on making sense of what it is like to be in a state with autonoetic consciousness. Alternatively, they might argue that, given the existence of explanatory projects with different methodological commitments in philosophy of mind, it is perhaps too strong to require that all attempts to account for autonoesis should satisfy the source criterion. While we agree that these are potentially promising ways to react to the discussion above, they have not been articulated explicitly in the existing literature. For this reason, we restrict ourselves to pointing to this issue as one that deserves more attention in future work.

4 | METACOGNITIVE ACCOUNTS

Metacognitive accounts attempt to explain autonoesis not in terms of the nature of episodic memory representations but rather in terms of metacognitive states that accompany them. These accounts appeal, in particular, to metacognitive feelings—where metacognitive feelings are understood as conscious affective states produced by subpersonal monitoring of the subject’s cognitive processes or capacities (see Arango-Muñoz, 2014; Moulin & Souchay, 2014)—that they take to be associated with episodic remembering.

4.1 | Types of metacognitive account

Two metacognitive accounts have been proposed. The first, developed by Dokic (2014a, 2014b), attempts to explain autonoesis in terms of an episodic feeling of knowing; the second, developed by Perrin and collaborators (Perrin, 2018; Perrin et al., 2020; Perrin & Sant'Anna, 2022), attempts to explain it in terms of the feeling of pastness.

4.1.1 | The episodic feeling of knowing

Episodic memory, on Dokic’s view, has a two-tiered structure, in that it involves two distinct components: a first-order representation of a past event and an “episodic feeling of knowing” (2014a: 413). The latter component is meant to explain why memory feels “first-hand” or why it seems “to originate directly from one’s past experience, excluding the essential involvement of either reasoning or testimony” (2014a: 415). On Dokic’s view, the episodic feeling of knowing results, like other metacognitive feelings, from the subpersonal monitoring of multiple cues, including information availability and the fluency of information processing. Two features of this account merit emphasis. First, the episodic feeling of knowing is meant to differ from the feeling of pastness, which has traditionally been invoked in descriptions of the phenomenology of episodic remembering. By Dokic’s lights, the feeling of pastness pertains to the temporal orientation of memory, whereas the episodic feeling of knowing pertains to the relationship between memory and past experience. Second, the episodic feeling of knowing is meant to differ from the more familiar feeling of knowing (Koriat, 2000). The feeling of knowing, as standardly understood, pertains to the subject’s ability to retrieve information. It thus can occur in connection both with episodic and with semantic memory. The episodic feeling of knowing, in contrast, pertains to the relationship between information that has already been retrieved and past experience: when
recalling the unfolding of a given event, one feels that one’s representation of that event originates in one’s experience of it. It thus can occur only in connection with episodic memory and might therefore in principle be the factor that distinguishes episodic from semantic memory.

4.1.2 | The feeling of pastness

A rival metacognitive account has recently been proposed by Perrin and collaborators (Perrin, 2018; Perrin et al., 2020; Perrin & Sant’Anna, 2022). Like Dokic, Perrin and collaborators treat autonoesis as a metacognitive feeling. They agree, in particular, that autonoesis results from the subpersonal monitoring of multiple cues associated with remembering. Unlike Dokic, they maintain that autonoesis is best characterized as a feeling of pastness rather than as a feeling of knowing. The primary difference between the feeling of pastness, as they understand it, and the episodic feeling of knowing, as Dokic conceives of it, concerns the kinds of cognitive processes they are thought to track. The episodic feeling of knowing, according to Perrin et al. (2020), evaluates a subject’s ability to engage in a given cognitive task (for example, whether the subject can retrieve a certain piece of information). The feeling of pastness, in contrast, evaluates a subject’s performance in a given cognitive task (for example, whether the subject has successfully retrieved a certain piece of information). Since, for Perrin et al. (2020), autonoesis is best described as an experience that is associated with actual retrieval of content, rather than with one’s ability to retrieve content, they argue that the feeling of pastness is better suited than is the episodic feeling of knowing to account for autonoesis.

In response, Dokic (2021) has argued that a key aspect of the phenomenology of episodic remembering is that it involves a sense that more information is available for retrieval, which can happen not only before but also after content is retrieved. When, for instance, one remembers one’s last meal in a restaurant and conjures up an image of oneself waiting to be seated, one’s memory is accompanied by a sense that more information is available for retrieval—information about the dish that one ordered, the layout of the restaurant, and so on. Since this sense pertains to one’s ability to engage in a cognitive task, Dokic suggests, it is not mistaken to think of the episodic feeling of knowing as contributing to the phenomenology of retrieved representations. It is unclear, however, how this feeling of availability of information relates to autonoesis.

4.2 | Assessing metacognitive accounts

These metacognitive accounts are situated similarly with respect to the objective and subjective episodicity questions and fare similarly with respect to the immediacy and source criteria.

4.2.1 | Objective and subjective episodicity

Metacognitive accounts are best seen as attempts to answer primarily the objective question. There are two different ways in which this question can be answered, which rest on different ways of interpreting the core claim made by metacognitive accounts. On the one hand, they might be viewed as claiming that what distinguishes episodic memory from semantic memory is, at the ontological level, the fact that only the former has a two-tiered structure, in which case they can be seen as providing a nonphenomenological response to the objective episodicity question. On the other hand, they might be viewed as claiming that what distinguishes episodic memory from semantic memory is, at the ontological level, the fact that only the former involves a certain metacognitive feeling, in which case they can be seen as providing a phenomenological response to the objective episodicity question—that is, that being an episodic memory just is being an autonoetic experiential state.

As with representational accounts, while it would seem like a natural next step for proponents of metacognitive accounts of the latter type to claim that subjects are able to identify episodic memories as being episodic in virtue of their having conscious experiences of a certain type—which would be explained by the involvement of metacognitive feelings—this is not a point that has been explicitly discussed in the context of those accounts, so it is not obvious whether they can be plausibly interpreted as answers to the subjective question. If, however, we assume that metacognitive accounts are indeed attempts to answer the subjective question, then metacognitive accounts seem to imply that subjects judge that they are remembering because their experiential states involve episodic feelings of knowing.
4.2.2 | Phenomenology and source

In treating episodic memory as having a two-tiered structure, metacognitive accounts are inspired by empirical metacognition research, in which it is standard to distinguish between first-order retrieval process and second-order metamemory processes (Metcalfe & Dunlosky, 2008), the latter being associated with autonoesis. In virtue of this aspect of the accounts, they seem to be particularly well positioned with respect to the source criterion. The same aspect of the accounts, however, puts them in a difficult position with respect to satisfying the immediacy criterion. The problem that metacognitive accounts face, given that they treat episodic memory as having a two-tiered structure, is that of explaining why we do not experience episodic memories as having such a structure. When one remembers episodically, one’s experience seems to be a single unit, in the sense that there is not a clear boundary between the features of the experience that are contributed by the retrieval process (the first-order representation) and the features that are contributed by metamemory processes (metacognitive feelings). Moreover, the information that memory originates in a past experience—a feature that metacognitive accounts attribute to metacognitive feelings—is arguably experienced as retrieved information, rather than as an addition to the information that is retrieved. All of this is in tension with the idea, central to empirical metamemory research, that metacognitive states are states of which one can become aware as being distinct from first-order processes, and, more importantly, states of which one can become aware even when the relevant first-order processes fail to occur. In the absence of an account of the “experiential unity” of episodic memory, metacognitive accounts will have trouble explaining the immediacy of autonoesis, for they seem to entail that awareness of an episodic memory as episodic results not simply from having that state but rather from becoming aware that the state is accompanied by a metacognitive state that is distinct from it.

Along the lines of the concerns discussed in relation to representational accounts, it might be that the problems faced by metacognitive accounts with regard to immediacy are simply due to their focus being elsewhere—namely, on providing an account that satisfies the source criterion. One strategy that partisans of metacognitive accounts may pursue here is to deny that metacognitive states are easily introspectable—they might, for instance, insist that metacognitive feelings lie in the “fringe of consciousness” (Mangan, 1993; Norman et al., 2010). Since, to our knowledge, no attempts along these lines have been made in connection with metacognitive accounts of autonoesis, it is still very much an open question whether metacognitive accounts can indeed satisfy the immediacy criterion.

5 | EPISTEMIC ACCOUNTS

Epistemic accounts attempt to explain autonoesis by treating episodic memory as a retained epistemic state, where an epistemic state is a state of knowledge or a state that puts one in a position to have knowledge.

5.1 | Types of epistemic account

Two distinct epistemic states—apprehension and knowledge-how—have been invoked in this context, resulting in two types of epistemic account.

5.1.1 | Retained apprehension

Perhaps the most influential epistemic account is that put forward by Martin (2001), according to whom autonoesis is explained by the fact that episodic memory is the preservation of a past instance of “apprehension.” Apprehension is, for Martin, characterized by a distinctive mode of awareness that is only possible when one perceptually interacts with events in the external world, such that those events are made present to consciousness—they are, to put it differently, made perceptually available to one. To remember an event is to bring a past instance of apprehension back to mind, so that what one previously apprehended is re-presented to one. Episodic memory thus differs from semantic memory.
in that only the former involves an experience of re-presenting a past apprehended event at the level of consciousness. When one remembers semantically, one is aware only of facts, and although such awareness might be linked to a past instance of apprehension, this link is not made manifest to the one at the level of consciousness—that is, it does not involve the re-presentation of what was previously apprehended. It is worth noting that, although Martin acknowledges that apprehension and knowledge are closely related—instances of apprehension are situations in which we find ourselves in a position to form knowledge—he denies that, strictly speaking, episodic memory involves retention of knowledge of past events.

5.1.2 | Retained knowledge-how

The idea that episodic memory involves the retention of a specific kind of knowledge has recently and independently been developed by Soteriou (2008) and Hoerl (2022). Both authors argue that episodic memory involves the retention of knowledge of what it was like to experience an event, which is to be understood in terms of the retention of an ability. For Soteriou, such an ability consists in one’s being able to “to do something that can put oneself in a state of knowledge whose propositional content is a distinctive kind of answer to the question, ‘what was it like to apprehend that event?”’ (2008: 481). For Hoerl, it consists in the ability to conjure up a mental image of a previously experienced event (2022: 20). Despite these potentially different ways of characterizing the ability in question, both views offer similar explanations of what makes episodic memory distinctive at the level of consciousness: while both semantic memory and episodic memory may involve retention of knowledge that such-and-such was the case, only episodic memory involves retention of knowledge of how one came to know that such-and-such was the case—more specifically, because the exercise of one’s retained ability allows one to experience now what it was like to experience the event in a particular way, one becomes aware that the information remembered was acquired by means of experiencing it in the past.

There has not been extensive engagement among proponents of epistemic accounts. As noted above, one clear point of disagreement among different epistemic accounts concerns the nature of the epistemic state with which episodic memory is identified. Deciding which of the existing accounts is more appropriate would seem to depend, at least in part, on how one thinks of the epistemic role that episodic memory plays in consciousness. If episodic memory is thought to be a source of knowledge, then explaining it in terms of apprehension seems more appropriate. But, if episodic memory is thought to be a form of knowledge rather than a source of knowledge, then the accounts developed by Hoerl and Soteriou are more attractive. It is unclear whether there are other points of disagreement among epistemic accounts; work focusing on the relationships among these accounts would be welcome.

5.2 | Assessing epistemic accounts

These epistemic accounts are situated similarly with respect to the objective and subjective episodicity questions and fare similarly with respect to the immediacy and source criteria.

5.2.1 | Objective and subjective episodicity

Like representational and metacognitive accounts, epistemic accounts seem to be best interpreted as attempts to answer primarily the objective question. Since they hold that what it means for an experiential state to be an episodic memory is for it to be an epistemic state of a certain type, and since having an epistemic state of that type is just what it is for an experiential state to have autoneosis, they seem to be committed to the claim that being an episodic memory just is being an experiential state with autoneosis.

Concerning the subjective question, while it would also seem like a natural next step for proponents of epistemic accounts to claim that subjects are able to identify episodic memories as being episodic in virtue of their having consciousness experiences of a certain type—which would be explained by the involvement of apprehension or knowledge-how states—this is not a point that has been explicitly discussed in the context of those accounts, so, once again, it is unclear the extent to which they can be plausibly interpreted as answers to the subjective question. However, if we interpret them as attempts to answer the subjective question, then epistemic accounts seem to imply that subjects judge that they
are remembering because their experiential states involve a phenomenology of re-presenting a past experience (Martin) or awareness of what it was like to have an experience (Soteriou and Hoerl).

### 5.2.2 Phenomenology and source

Epistemic accounts are well positioned with respect to the immediacy criterion. Since, if such accounts are right, all it takes for one to have a memory with autonoesis is for one to have retained a specific type of knowledge, no additional step is required at the level of consciousness for one to become aware of an episodic memory as an episodic memory. The accounts thus provide a straightforward explanation of the immediacy of autonoesis. Like proponents of content-based representational accounts, however, proponents of epistemic accounts have said very little about the source of autonoesis. One possibility is that the question of the source of autonoesis is simply not within the scope of such accounts. If this is the case, then proponents of epistemic accounts are perhaps best interpreted as being engaged in an explanatory project other than that in which proponents of metacognitive and (attitude-based) representational accounts are engaged. Another possibility is that the question of the source of autonoesis is simply not within the scope of the accounts and that their proponents simply have yet to take that question up. If this is the case, however, then proponents of epistemic accounts appear to face an important explanatory challenge.

The challenge arises because epistemic accounts, in treating episodic memory as an epistemic state, assign a central explanatory role to normative or evaluative terms, such as apprehension and knowledge-how. Such terms are traditionally thought to have no place in purely descriptive explanations of the sort offered by the empirical sciences. The empirically respectableability of epistemic accounts is thus questionable. To see why, note that epistemic accounts in effect assume that remembering is a success state. They assume, in particular, that the occurrence of remembering presupposes, among other things, that the rememberer experienced the remembered event when it occurred and that his
current memory of it is due to that past experience. This assumption is difficult to reconcile with the fact that we appear to have experiences of the same type as the experiences that we have when remembering successfully when we entertain memories that, unbeknownst to us, are false or based on second-hand information and that those experiences appear to be produced by the very same processes that are at work when we remember successfully.\(^\text{16}\) In light of this fact, an epistemic account of the source of autonoesis would need to explain how the mechanism responsible for the production of autonoesis manages to keep track of the success of occurrences of remembering, allowing the difference between successful and unsuccessful remembering to show up in their respective phenomenologies. It is by no means obvious that such a mechanism exists or even what it might in principle look like. As far as the source criterion is concerned, then, the use by epistemic accounts of normative terms saddles them with a particularly heavy explanatory burden.

6 | CONCLUSIONS AND FUTURE DIRECTIONS

To summarize, all three accounts are plausibly viewed as targeting objective episodicity, but it is not clear whether they should also be viewed as attempts to make sense of subjective episodicity. Representational and epistemic accounts seem to do justice to the immediacy of autonoesis but—with the exception of attitude-based representational accounts—struggle to explain the source of that phenomenology, whereas metacognitive accounts seem to be able to explain the source of autonoesis but may fail to do justice to its immediacy (Tables 1 and 2).

There is, of course, much more to be said about how the existing accounts of autonoesis relate to one another and also about their prospects for answering the objective and subjective questions. It was not, however, our goal here to provide a detailed analysis of these issues. Rather, as noted in Section 2.4, given the fact that there has been little engagement among existing accounts of autonoesis, our efforts were directed at providing an overview of these accounts and providing a minimal framework to set the stage for future work in the area.

Overall, our discussion highlights two ways in which the philosophical literature on autonoesis is underdeveloped and in which it would benefit from further work. First, additional clarity is required regarding the role played by the concept of autonoesis in philosophical theorizing about memory. Should accounts of autonoesis aim to explain how subjects distinguish between episodic and semantic memory (responding to the subjective episodicity question), or should they aim to explain the very nature of the difference between episodic and semantic memory (responding to the objective episodicity question)? These two questions are rarely distinguished in the literature, and while the assumption seems to be that answering the objective question naturally gives us an answer to the subjective question, it is not obvious whether that assumption is warranted. Second, additional clarity is required regarding the criteria that accounts of autonoesis are expected to meet. Does the acceptability of an account depend on whether it meets both of the criteria employed here? Are there other criteria, such as explanatory consilience between existing philosophical and psychological perspectives, that should be considered? Answering these questions will enable us not only to determine whether and to what extent existing accounts are in competition with one another but also to establish a consensus on the standards against which competing accounts should be evaluated.

AUTHOR CONTRIBUTIONS

André Sant’Anna: Conceptualization (equal); investigation (lead); writing – original draft (lead); writing – review and editing (equal). Kourken Michaelian: Conceptualization (equal); investigation (supporting); writing – original draft (supporting); writing – review and editing (equal). Nikola Andonovski: Investigation (supporting); writing – review and editing (equal).

FUNDING INFORMATION

André Sant’Anna is funded by an Alexander von Humboldt Postdoctoral Fellowship. Nikola Andonovski’s research is funded by the European Union’s Horizon research and innovation program under the Marie Sklodowska-Curie grant agreement 101062754.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.
DATA AVAILABILITY STATEMENT
Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID
André Sant’Anna  https://orcid.org/0000-0002-2239-7243
Kourken Michaelian  https://orcid.org/0000-0002-5205-3046
Nikola Andonovski  https://orcid.org/0000-0001-6957-0261

RELATED WIREs ARTICLES
What memory is
Opening the doors of memory: Is declarative memory a natural kind?

ENDNOTES
1 Tulving sometimes (e.g., 2002) referred to chronesthesia in addition to autonoesis, describing the latter as a form of consciousness of the self and the former as a form of consciousness of time. But he more often treated autonoesis as having a temporal dimension, describing it, for example, as “the neurocognitive capability of normal adults to become aware of their existence in subjectively experienced time” (2001: 23). In practice, the difference between autonoesis and chronesthesia has usually been a matter merely of emphasis, with both the self and time figuring in both forms of consciousness (Szpunar, 2011). Chronesthesia will accordingly be set aside in what follows.

2 While the article treats the accounts in question as pertaining to episodicity in particular, their proponents have sometimes framed them in more general terms, suggesting that they identify a feature that marks episodic memory off not only from semantic and other forms of memory but also from nonmnemonic mental states such as imagination. Some of the accounts reviewed below are thus framed by their authors as pertaining not only to the episodicity question but also to the mnemicity question (Michaelian & Sutton, 2017): what distinguishes between memory, on the one hand, and imagination and related mental states, on the other? The article’s focus on episodicity rather than mnemicity is not meant to suggest that the former is more basic than the latter but simply to encourage more systematic work on an issue that has so far been approached only in a piecemeal fashion. For recent work on mnemicity, see Byrne (2010); Michaelian (2016); Sant’Anna and Michaelian (2019); Sant’Anna (2020); Barkasi and Rosen (2020); Kind (2021); Fish (2021); Mahr et al. (2023); Rivadulla-Duró (2022).

3 See, however, Section 2.3 below, where we distinguish between two different ways of interpreting the episodicity question.

4 One theoretical possibility, which could be motivated by a more careful consideration of the question whether semantic memory has the feature responsible for autonoesis, is that the difference between episodic memory and semantic might be one of degrees. More specifically, it might be that, rather than being absent in semantic remembering, autonoesis is only present in weak or very weak forms in semantic remembering. What the implications of taking autonoesis to be a graded form of consciousness would be for existing accounts is an interesting question for future research.


6 In more recent work, Dokic (2014a, 2021) has defended a metacognitive account of autonoesis. See Section 4.

7 Fernández (2019) argues that features of the phenomenology of episodic remembering including the feeling of pastness, the feeling of ownership, and mental time travel can be explained in terms of episodic memory’s self-referential content. It is unclear, however, whether he means to suggest that these feelings are distinct from each other or whether he instead takes them to boil down to the same thing. Determining whether there is more than one feeling at issue here would require providing a relatively precise description of the relationship between content and phenomenology, which Fernández has not done so far. See Perrin and Sant’Anna (2022) for further discussion.

8 Despite taking autonoesis to be an outcome of a metarepresentational capacity, Mahr and Csibra do not appear to be committed to the claim that episodic memories understood as experiential states themselves have a metarepresentational structure. Although they are not explicit on this point, the fact that they are willing to characterize autonoesis in terms of an “attitude” and the fact that they speak of autonoesis as being the outcome of a metarepresentational capacity as opposed to being a metarepresentational state both speak in favor of grouping their account with the representational accounts discussed in this section rather than with the metacognitive accounts discussed in Section 4. See Pan (2022) for further discussion.
Metacognitive feelings are sometimes referred to as “epistemic” feelings. Dokic, whose metacognitive account of autonoesis is discussed in this section, has also referred to them as “noetic” feelings (Dokic, 2012, 2014b). Dokic’s use of the term should not be confused with Tulving’s: Dokic’s concept of a noetic feeling applies not only to noesis in Tulving’s sense, but also to autonoesis and to other feelings potentially involved in remembering, as well as to a variety of metacognitive feelings involved in cognitive processes other than remembering.

Though see Dokic (2021), where Dokic argues that the episodic feeling of knowing also concerns the subject’s capacity to retrieve more information in episodic memory (see also Section 4.1.2 for discussion). One concern with this claim is that it is unclear how, in light of this new characterization, the episodic feeling of knowing is to be distinguished from the semantic feeling of knowing.

See also Hoerl (2022: 8–9), who points out that the existence of the metacognitive feelings supposed to be involved in episodic remembering are not clearly revealed to introspection.

Note that this point pertains specifically to metacognitive accounts of autonoesis, which, unlike accounts that merely accept the causal involvement of metacognition in retrieval—for example, Mahr and Csibra’s (2018) representational account—claim that metacognitive states are constitutive of episodic memories understood as experiential states.

The contrast here is with cases, such as hallucination, in which it merely seems to one that one is perceptually interacting with events in the external world. These cases do not involve apprehension in the way Martin uses the term, and hence do not genuinely make events perceptually present to one. In the background here is a disjunctivist view of perception that denies that veridical and hallucinatory experiences are experiential states of the same kind (Martin, 2004; see Soteriou, 2020 for an overview).

The relationship between the epistemic account articulated in Hoerl (2022) and the representational account defended in Hoerl (2001) (see Section 3.1.1) is an open question. Given the programmatic nature of his proposal, Hoerl (2022) is quite explicit regarding the fact that there is still much work to be done to flesh out the nature of the knowledge that is retained in memory. It seems, however, that one natural way of doing so would be to say that the ability retained in memory is that of conjuring up a mental image that represents the spatial enabling conditions of the subject’s previous experience. Whether the result would amount to a representational or to an epistemic account of autonoesis is a further open question.

For a detailed discussion of the relationship between normative and descriptive or empirical notions of remembering, see Craver (2020).

Proponents of the epistemic account might respond here by saying that this begs the question against them. Even if that is correct, the original concern remains: the basic explanatory project that motivates epistemic theories may simply be incompatible with the aim of providing an account of the source of autonoesis.

**FURTHER READING**


**REFERENCES**


How to cite this article: Sant’Anna, A., Michaelian, K., & Andonovski, N. (2023). Autonoesis and episodicity: Perspectives from philosophy of memory. WIREs Cognitive Science, e1665. https://doi.org/10.1002/wcs.1665