



Memory-based reference and immunity to error through misidentification

Manuel García-Carpintero¹

Received: 3 October 2023 / Accepted: 1 June 2024
© The Author(s) 2024

Abstract

Wittgenstein distinguished between two uses of ‘I’, one “as object” and the other “as subject”, a distinction that Shoemaker elucidated in terms of a notion of *immunity to error through misidentification* (‘IEM’); in their use “as subject”, first-personal claims are IEM, but not in their use “as object”. Shoemaker argued that memory judgments based on “personal”, *episodic* memory are only de facto IEM, not strictly speaking IEM, while Gareth Evans disputed it. In the past two decades research on memory has produced very significant results, which have changed the philosophical landscape. As part of it, several new arguments have been made for and against the IEM of personal memories. The paper aims to defend the Shoemaker line by critically engaging with some compelling recent contributions.

Keywords Memory · First-personal reference · *De se* attitudes · Self-knowledge · Immunity to error through misidentification

1 Preamble: immunity to error through misidentification and self-reference

Wittgenstein (1958, pp. 66–67) distinguished between two uses of ‘I’, one “as object” and the other “as subject”, a distinction that Shoemaker (1968) elucidated in terms of a notion of *immunity to error through misidentification* (‘IEM’ henceforth, also for *immune* ...). Shoemaker (1968, p. 557) specifies the sort of mistake that IEM excludes as follows: “to say that a statement ‘*a* is φ ’ is subject to error through misidentification relative to the term ‘*a*’ means that the following is possible: the speaker knows some particular thing to be φ , but makes the mistake of asserting ‘*a* is φ ’ because, and only because, he mistakenly thinks that the thing he knows to be φ is what *a* refers to”. Shoemaker (1970, p. 270) offers a memory illustration: “if I claim on the strength of

✉ Manuel García-Carpintero
m.garciacarpintero@ub.es

¹ LOGOS/BIAP-Departament de Filosofia, Universitat de Barcelona, Barcelona, Spain

memory that I saw John yesterday, and have a full and accurate memory of the incident, it cannot be the case that I remember someone seeing John but have misidentified that person as myself; my memory claim ‘I saw John’ is subject to error through misidentification with respect to the term ‘John’ (for it could have been John’s twin or double that I saw), but not with respect to ‘I’”. The error can also affect ‘I’ if, e.g., I base the claim on a photograph of what I take to be me looking at John. I may know on that basis that someone who looks like me sees John but misidentify him as me. This would be a contrasting use of ‘I’ “as object”.

Shoemaker (1970) also distinguishes two kinds of IEM depending on how strongly we read the modals (‘possible’, ‘cannot’, ‘could’) in his characterizations. If it is conditions in the actual and nearby worlds that rule out the possibility of the specified mistake, then the judgment is *de facto* IEM relative to evidence E. If the conditions in every logically possible world rule out the mistake, then the judgment is *logically* IEM relative to E. He argued that memory-based claims like ‘I saw John yesterday’ are only *de facto* IEM, by considering science fiction cases involving brain-splitting and brain-transplants. Evans (1982) contested his views.

Recent research on memory points to its “constructive” nature, manifested by the *observer* or *third-person vs. field* or *first-person* character of some of our memories (Rice, 2010). In *field* memories we remember ourselves occupying the visual perspective from which we perceived the remembered scene. In *observer* memories, we recall ourselves as a participant in that scene. Several new arguments have been made on this basis for and against the IEM of personal memories, examining more ordinary cases than earlier discussions. My goal here is to defend the Shoemaker line that memory claims are at most *de facto* IEM (Sect. 3); I’ll address (Sect. 4) arguments for Evans’s view by Fernández (2021), Hamilton (2007), Lin (2020), and Perrin and McCarroll (2023). I’ll provide first background on IEM (Sect. 1) and on episodic memory (Sect. 2). I’ll explain in Sect. 1 how I think Shoemaker’s notion of IEM should be understood. I’ll offer a model on which IEM is hypothesized to be a psychological natural kind, so that there is right and wrong about this; my points in Sect. 4 depend in part on the view that some arguments that memory is (logically) IEM assume incorrect accounts of the phenomenon.¹

Shoemaker defines IEM for linguistic acts, while the phenomenon I am interested in concerns mental attitudes, memory judgments and the memory episodes on which they may be based. Here I’ll understand judgments as inner assertions—assertions in inner speech, on the assumption that inner speech, at least in some cases, is actual speech as opposed to imagined speech (Gregory, 2016; Kompa, 2023). Shoemaker’s characterization thus extends to thoughts. IEM is a phenomenon discerned by philosophers; however, debates whether some attitudes (self-ascriptive judgments based on episodic memories in our case—or on introspection in a related debate, García-Carpintero (2024)—are IEM presuppose that it is not a fabricated construct whose character can be stipulated as we wish; there are right and wrong accounts of the phenomenon to be explained, and of course also of how to explain it.

¹ I’ll skip the *logically* vs. *de facto* qualification when context makes it clear which is meant.

I suggest that we think of this on the well-known Kripke-Putnam model for natural kinds like *water*. We rely on superficial criteria to classify a stuff as *water*: thirst-quenching transparent, odorless, tasteless liquid filling rivers, lakes and oceans. But these are just helpful criteria to identify paradigms. They do not specify what it is to be water, what is really common to all instances of the stuff; something that is not water might share them. It is rather a non-manifest “essence” (*being constituted by H₂O molecules*), whose instantiation explains, among other facts, the superficial reference-fixing features. Similarly, I’ll assume that the debates about the extent of IEM presuppose that the features that Shoemaker mentions distinguish central cases of a psychologically important kind with an explanatory real nature.

Seeger (2015) allows us to streamline Shoemaker’s characterization, encompassing Pryor’s (1999) cases of wh-misidentification, see below. We first define *error through misidentification*, ‘EM’ henceforth. The *source object* in a judgment of the form *a is F* is the object, if any, from which the predication content F derives—the one that epistemically grounds it. The *target object* is *the object a to which that content is applied, if any*. The definition of EM is then this:

EM A judgment *a is F* is in *error through misidentification* iff the source object is different from the target object.²

The core idea is that in *EM* cases we can tease apart, and retain, the justification for the existential claim, *that something is F*; it is independent from the justification for the singular claim, *that a is F*, which might be false or unjustified.³ In IEM cases, in contrast, the evidence gives us a “package deal”—indicating at once that something *is F*, and that *it is a which is F* together with it. Shoemaker states and illustrates the “package deal” metaphor thus: “[...] in being aware that one feels pain one is, tautologically, aware, not simply that the attribute *feel(s) pain* is instantiated, but that it is instantiated *in oneself*” (Shoemaker, 1968, p. 563f.; his emphases). If a judgment made on grounds E is liable to EM, I will say that it is *vulnerable to error through misidentification* (‘VEM’ henceforth, also for *vulnerability ...*), relative to E; otherwise, it is IEM relative to E. This can be so *logically*, in all worlds, or *de facto*, just in nearby worlds.

This is thus what I take to be the initial, manifest feature of IEM attitudes, but what explains it, on the *natural kind* model? Thinking of paradigm instances like ‘I am in pain’ or ‘I saw John’ based on introspection, and non-instances like the latter judgment based on a photograph of what I take to be me looking at John, Anscombe (1975) worried that we might be led to explain IEM on the view that we are Cartesian egos. We know what we really are when we access our features as thinking beings (*being in pain, seeming to remember John*); this is why in such cases we get the “package deal” and misidentification is not possible, unlike when we self-ascribe properties that we come to know the way we know material beings. She advanced an alternative deflationary explanation: in IEM judgments ‘I’ doesn’t refer; it works like ‘it’ in ‘it

² This slightly modifies Seeger’s (2015, p. 2) proposal. McGlynn (2021, p. 2305) has a more precise definition developing this core idea, cf. Morgan (2019, pp. 446–447) and Palmira (2020, Sect. 2) for related characterizations and McGlynn (2021, pp. 2309–2310) for discussion.

³ This generalizes what Pryor (1999, p. 296) says about memory. Among others, Wright (1998, p. 19), Campbell (1997, p. 89) and Prosser (2012, pp. 161–162) share this core notion.

rains'. There cannot be a misidentification of the described sort not because 'I' refers to a Cartesian ego, but rather because there is no identification in the first place. Relying on ideas from David Lewis, Recanati (2007, 2009, 2012) develops a similar account of the IEM of "basic" or "implicit" *de se* thoughts. The Anscombe-Recanati view is a particular version of Evans's (1982, pp. 180–181) *Simple Account* of IEM, on which VEM judgments depend on an "identification" of the referent while IEM judgments are identification-free. Evans's version avoids Anscombe's problematic commitment to the non-referential character of 'I'. He is however willing to extend logical IEM to first-personal claims with more "external" reaches (memory, proprioception) to resist like her Cartesian accounts (cf. Fernández, 2014, p. 376; 2021, p. 644).

I have argued however (García-Carpintero, 2015, 2016, 2017, 2018) that the Simple Account in all forms fails to explain IEM. A main reason against the no-reference version is that there are examples involving indexicals other than 'I' for which the view that they don't refer is even more implausible than nonetheless appear to be IEM: 'you are standing very close', 'he is a long way off', Wright (2012); 'this keyboard is black', Peacocke (2008), all based on perception.⁴ These cases also provide reasons against other versions of the the Simple Account because they don't appear to be "identification-free" but assume an "identification" of the referents.⁵ Also, there are judgments that don't appear to rely on any identification—Pryor's, 1999 cases of *wh-misidentification*—but appear to be VEM, as defined above. These are cases in which the subject makes a justified existential judgment (*someone drank my beer*), without committing to there being a unique witness, and then goes on to misidentify one (*Alex drank my beer*).

I think that what Palmira (2020) and Coliva and Palmira (2024) call *Metasemantic account* (García-Carpintero (2015, 2016, 2017, 2018), Palmira (2020, 2022), Verdejo (2021)) offers a much better explanation, without surrender to a problematic Cartesian ontology. For present purposes we may do with the guiding idea. Consider *Julius invented the zip*, assuming that the referent of 'Julius' is stipulated to be *the inventor of the zip*. This looks IEM on the previous characterization of the phenomenon; indeed, *logically* IEM: if in fact somebody invented the zip, it cannot be other than Julius. In this case, *being the inventor of the zip* plays a metasemantic role in fixing which *Julius* is at stake. Similarly, in intuitively IEM demonstrative examples ('you are standing very close', 'he is a long way off', 'this keyboard is black'), material derived from perception that "entails" the ascribed features (*standing close*, *being a long way off*, *being black*) helps fixing the referent. On the metasemantic view IEM is a semantic phenomenon involving epistemically significant reference-fixing evidence.⁶

⁴ To elaborate: 'this keyboard is black' appears to be IEM on our definition above because, if the existential "part" of the judgment is true and justified (i.e., something in the immediate vicinity of the speaker, the *source*, is black), it doesn't seem to be possible that it is not the referent of 'this keyboard'. The intuitive reason—which the Metasemantic account I outline below develops—is that the referent of the demonstrative 'this keyboard' is in part perceptually fixed in the relevant context as the *black* object before the speaker.

⁵ I am skipping over technically complicated epistemological and semantic issues for a wider readership, in particular on whether "identifications" are premises in tacit inferences or rather background presuppositions; cf. Coliva and Palmira (2024), and references above.

⁶ Campbell (1997, p. 70) argues against a view like the one just outlined with the following example: "if you judge, 'that chair is yellow', it may be that you thereby know of something that it is yellow, but

The methodological claims I have made about IEM play a crucial role in my argument for Shoemaker's line about memory. To sum up, I assume that the phenomenon, here identified by the Shoemakerian characterization above, aims to discern a psychological natural kind. There are thus better and worse specifications of the phenomenon to be explained. Some theoretical proposals like the no-reference and the more general Simple Accounts fail to explain it and hence to properly define the kind; the Metasemantic account offers better prospects. In the next section I'll advance similar methodological assumptions about personal memory.

2 Episodic memory and the imagination

Our debate concerns *episodic* or *personal* memories. I'll borrow from Robins (2020) an intuitively apt characterization: "Episodic memory differs from other forms of memory not only in its content—being focused on personally-experienced events rather than general facts or skills, as is the case for semantic and procedural memory, respectively—but in how that content is presented to the rememberer. Episodic memories are first-personal and represented through autoegetic consciousness, allowing the rememberer to (at least sometimes) feel as if they are mentally time traveling back to the event or experience being recalled" (Robins, 2020, pp. 470–471).

As I have characterized debates about IEM in the previous sections, they primarily concern judgments, relative to their epistemic grounds; in our case, judgments based on episodic memories. However, in the recent literature I'll be critically examining in the following sections, the debates sometimes primarily target episodic memories themselves.⁷ This presupposes controversial views about the nature of such memories,⁸ echoing similar controversies about perception that are also relevant here because, as the characterization by Robins just quoted shows, intuitively episodic memories target earlier perceptual experiences of the same subject. The controversial assumptions are, first, that mnemonic experiences and the perceptual episodes from which they derive *are contentful representations*; and second, that their contents are distinctively

Footnote 6 continued

that thing is not the chair, if, for instance, the chair is transparent and set against a yellow background. If your judgment is mistaken, you can rectify it by retreating to the more cautious, 'At any rate, something is yellow'." García-Carpintero (2018) retorts that this just shows that the IEM in these cases is merely *de facto*.

⁷ Thanks to Denis Perrin and André Sant'Anna for pointing this out. My assumption is standard in philosophical discussions of the IEM of memory or perception; cf. for instance Pryor (1999), Hamilton (2007, p. 411), Fernández (2014, 2021). Given that it is judgments based on the relevant episodic memories that we target, nothing of substance is lost, I think. We thus avert the controversial issues to be mentioned in a moment. Another important reason to conduct the discussion in the standard way is that perceptual experiences are not usually considered apt for epistemic evaluation, and the same seems to apply to mnemonic seemings or episodic memories (Senor, 2022); but accounts of IEM presuppose epistemic evaluability.

⁸ Henceforth I'll omit 'episodic', but it should be understood unless otherwise indicated. Like Robins (2020), and given the representational view to be advanced shortly, I'll avoid the circumlocutions required by the factivity of 'perception' and 'memory' by focusing just on perceptual and mnemonic *experiences* or *seemings*, even when this is not explicitly signaled. Shoemaker (1970) introduced his technical notion of q(uasi)-memories for similar reasons. I discuss mnemonic experiences because I think any good account of memory will need to posit them as constituents of memories; q-memories as defined by Shoemaker may just be a theoretical posit tailored to the debate about IEM, without any other explanatory role.

first-personal, *de se*. Naïve realists or disjunctivists (cf. Soteriou, 2020) reject the first assumption for perception; Martin (2001) extends the view to episodic memories. These views forgo the notion that perceptions and memories are representations; they think of them instead as genuine relations of acquaintance with external items, objects, events or facts. They thus take them to have different natures than perceptual experiences or apparent memories that “feel identical from the inside” but misrepresent target situations. There are serious objections to disjunctivist views for perception and memory on account of this (cf. e.g. Mehta, 2023; Robins, 2020, respectively) that allow me to grant the first assumption—albeit I suspect my arguments might be run in disjunctivist-friendly terms. I will therefore think of perceptual experiences and apparent memories as intentional contentful representations that may represent incorrectly.

What about the second assumption? What would it be for perceptual and mnemonic experiences to have first-personal, *de se* content? On the *metasemantic* view referenced in Sect. 1, one main aspect of a *de se* representation is that its referent is fixed by a self-reference rule, understood token-reflexively: the representation is assumed to pick out the thinker of the token-attitude of which it is part. The metasemantics for the relevant representations assumes richer descriptive features to zero in on the referent. They include dispositions to “coordinate” (Fine, 2007) or “put together” the information predicated of the referent with information predicated through other referential items, by “placing them” in the same “file” (Recanati, 2012). In the case of the first-person, relying on Shoemaker Peacocke (2014, pp. 86–99, 106–113) spells out the most significant dispositions. Roughly, they present the self as the object whose body is the one from which that self normally perceives the world, receives bodily information, and executes basic actions. Such dispositions to treat information as concerning the same object can be associated not just with ‘I’-representations in judgments, but also with related representations in perceptual and mnemonic experiences. I’ll thus also grant the second assumption, spelled out along these lines.

Let’s move now to the continuum-discontinuum debate on “mental time travel”, cf. again Robins (2020). Episodic memories intuitively aim to put us in touch with our past; we also consider and plan for our future by means of the imagination, among other things by exploring possibilities in decision-making. The imagination can also be oriented towards the past, for instance when we represent with regret what might have been but wasn’t. Intuitively, perceptual and mnemonic experiences, and imaginings, can be distinguished relative to different dimensions, which can be enlisted by philosophical theories to characterize attitudes. We cannot sometimes be certain on phenomenological grounds whether we are perceiving, remembering, or imagining—think of a tune that tends to come unbidden, or bells that regularly toll in a nearby church. But standardly we do immediately distinguish those attitudes on phenomenal grounds.⁹ Functionally, we take perceptual experiences in normal conditions to be caused by a

⁹ I mostly agree with Teroni’s (2024) account of how memories (mnemonic experiences really) are identified by the subject, which assumes as I do that this is a “personal-level” (as opposed to *subpersonal*) matter. It allows for the mistakes just mentioned both of omission (failing to realize that an attitude is a mnemonic experience) and commission (mistaking as one something that is not, perhaps an imagining), but it assumes that typically they are reliable. I also agree with him that, given sensible constraints he sets up, accounts like Fernández’s (2024) in terms of the contents of memories are inadequate; adequate accounts should

typically just occurrent instance of the represented condition, which goes on existing concurrently with the experience. We take mnemonic experiences to be caused in normal conditions by a previous perception of the condition we take ourselves to remember, which hence occurred earlier. And we take imaginings to be causally unconstrained in any of those ways.¹⁰ Both perceptual and mnemonic experiences may be “hallucinatory”, and hence, like imaginings, they may in fact lack causal connections with the situations they represent. But the described features still distinguish the attitudes *normatively*: whether they really meet their “quasi-simultaneous” or “past” causal conditions, perceptual and mnemonic experiences would differ from imaginings by being beholden to norms requiring that those conditions be met.¹¹ Imaginings are not beholden to those causal norms, but apt alternative norms can also be articulated for them (Hunt, 2023). For one thing, they should be sufficiently rich and detailed to appropriately serve the projects to which they are put.

Like Robins (2020), I’ll assume a version of *discontinuism* like the one just outlined based on the traditional attitude (force) versus content distinction, which is similar to hers. Roughly, continuists contend that mnemonic representations don’t differ constitutively from corresponding representations provided by the imagination (cf. Michaelian et al., 2020), while discontinuists claim that they do—in my view on account of the normative causal conditions just outlined. The extent to which episodic memory is more reconstructive than preservative—as research in the past two decades reveals and the case of observer memories illustrates—and the overlap in the neurological bases of memories and imaginings appears to support continuism. But I take it that Robins (2020), among others, offers good reasons to uphold some form of discontinuism compatible with those results; cf. also Michaelian and Robins (2018) for ways that a causal condition on proper memories can be articulated, and the challenges they confront.¹²

To apply the force/content distinction to the attitudes we are concerned with here presupposes that they may share contents, and hence (on standard assumptions about the relevant imaginings) that such contents are propositional. This is perfectly in sync with current research in semantics, which takes contents to be highly abstract: it is not just utterances in linguistically very different languages like Basque and English that may share contents, but also representations in different media, like the depictive representations of perceptual experiences and the “propositional” representations of

Footnote 9 continued

rely instead on traits of the attitude type. Teroni offers a compelling one on which it features a “feeling of familiarity”; cp. Dokic (2014, Perrin and Sant’Anna (2022).

¹⁰ Fernández (2017) offers a clear articulation of the intuitions described here, cf. also Gregory (2023), Mehta (2023).

¹¹ In the process of articulating the disjunctivist view of perception and memory we have questioned, Martin (2001) nicely describes the web of causal relations between perceptual experiences, mnemonic experiences, and the external events they target. He describes the phenomenology of these attitudes normatively: “the picture of experiential memory sketched here is *an ideal* to which we can only sometimes conform, but which nevertheless informs centrally one aspect of why we value episodic memories in the way that we do” (Martin 2001, p. 282). Properly articulated, the normative account I envisage should capture the sense in which the obtaining of the relevant causal relations is “an ideal” for the experiences.

¹² Disclaiming preservativism makes it mandatory to articulate accuracy criteria for correct memories, which is not an easy task. Dings et al. (2023) offer some ideas, and the plausible suggestion that there might be different criteria for different projects that memories may serve.

judgments in inner speech. We can think of those contents as classes of possible worlds or, better, “smaller” worlds like conditions or situations (Kratzer, 2021).¹³ Differences in inferential potential among attitudes with the same abstract content (like those Fregeans invoke to posit modes of presentation) can on this view be explained in terms of differences in the content-vehicle. On the token-reflexive view, subjects of conscious attitudes are consciously aware not only of their contents, but also of features of their vehicles. Imaginings may feature images; however, propositional imaginings may come without them (when imagining, say, that the economy is improving), and attitudes other than imaginings may come with images (say, depictive assertions). Most attitudes are hybrids of imagistic features and “propositional” features as these are traditionally understood. It is a philosophical challenge to properly characterize their contents, but it is one more we don’t need to deal with here.

I’ll conclude this presentation of background assumptions on memory with the field vs. observer distinction. In a review of the philosophical aspects of the debate about observer memories, McCarroll and Sutton (2017) describe the difference in this way:

The imagery involved in remembering past episodes in one’s life often involves visual points of view. When we recall a past event, we usually adopt the same perspective that we had at the time of the original experience. We see the scene as we originally saw it from a first-person or ‘field’ perspective. Sometimes, however, we recall the past event from an external visual perspective, from a position we didn’t occupy at the time of the original episode. In such cases, we view ourselves in the remembered scene, as from a third-person ‘observer’ perspective (McCarroll & Sutton, 2017, 114)

Rice (2010) offers a helpful review of the empirical literature. For some reason, although we can easily tell by introspection that some of our memories are “observer” (the point had been noticed by many researchers earlier, including Freud), noticing it feels like a revelation; at least this was so in my case. I have a vivid apparent memory from the time my sister was born; it is unusual, because she is only one year and a half younger than I am. The pregnancy had been difficult, and I had spent the final few weeks with a childless, child-loving aunt of mine. I was then taken to meet my sister. In the field version of this memory (the first I remember having), I am at the door of a bedroom; an adult (my father, I was told by my parents, who confirmed the spatial layout that I described) is in the entrance at my left; my mother and my newborn sister are in the bed, perpendicular to the direction of my sight, headboard along the wall at my left; I feel an overriding wave of intense hate towards the intruder. Although initially this was a field memory, as described, usually when I call it back now it comes as an observer memory, and then (as the empirical research suggests is common) the emotional aspects are subdued or simply absent. I can now easily switch

¹³ Rowlands (2018) and Sant’Anna (2018) argue that memories don’t have propositions as intentional objects, but this is because they assume a conception of propositions as linguistically structured; as indicated in the text, I assume that the “non-conceptual” Fregean aspects of the relevant memories that, as they correctly point out, apt characterizations should posit, can be captured as features of their vehicles. Fernández (2017) assumes a view of contents similar to mine. However, he has the *perspective* and *causal* conditions intuitively distinguishing perceptual experiences, mnemonic experiences and imaginings as part of their contents, while I think they characterize instead their distinctive forces—cf. Recanati (2007), Robins (2020).

at will from the field to the observer perspective. This is to be expected given the reconstructive character of memory; as Vendler (1979) shows, we can easily switch from one perspective to the other in our imagination:

We are looking down upon the ocean from a cliff. The water is rough and cold, yet there are some swimmers riding the waves. “Just imagine swimming in that water” says my friend, and I know what to do. “Brrr!” I say as I imagine the cold, the salty taste, the tug of the current, and so forth. Had he said “Just imagine yourself swimming in that water”, I could comply in another way too: by picturing myself being tossed about, a scrawny body bobbing up and down in the foamy waste. In this case, I do not have to leave the cliff in the imagination: I may see myself, if I so choose, from the very same perspective. Not so in the previous case: if I indeed imagine being in the water, then I may see the cliff above me, but not myself from it (Vendler, 1979, p. 161)

The first imagining, which Vendler calls “subjective”, targets a “field” experience, involving visual aspects (the perspective from the sea of the top of the cliff) and other sensory experiences (thermoceptive, kinesthetic, ...), all “from the inside”. The second induces an “observer” imagining (which Vendler calls “objective”). Core imaginings are mental actions, and thus we can switch the perspective in response to instructions, at will; but memories are not similarly under our control. *Prima facie* however, cases like the previous example suggest that we can also enjoy memories of both kinds, targeting the same situation.

In the quoted text, Vendler suggests a dubious assumption that anticipates both his analysis afterwards of the “objective” reading, and the dismissal of observer memories he later explicitly articulates. As we will see, it is shared by other researchers; but in my view it is unwarranted.¹⁴ This will play a crucial role in our discussion. The analysis Vendler provides for the objective reading has it that “imagin[ing] yourself swimming in the water” is equivalent to “imagin[ing] seeing yourself swimming in the water” (Vendler, 1979, p. 164). As D’Ambrosio and Stoljar (2021) show, this should be rejected. We don’t need to go into the compelling semantic analysis they provide, which supports their rejection. The point we need is intuitive enough to be made without articulating a semantics validating it. There is no need to assume, as Vendler’s analysis does, that the perspective on the cliff from which I imagine me being “seen” swimming is also mine; in fact, it doesn’t need to be that of any conscious being, but just the “perspective” of a camera. Through TV, films, and paintings we are familiar with imagining alien perspectives, vacated (the “camera eye”) or assumed to the POV of some character: it doesn’t take much to realize that most of Robert Montgomery’s 1947 *Lady in the Lake* is presented as the visual perspective of the main character, private eye Philip Marlowe.¹⁵

As indicated, Vendler considers the issue whether the “objective” (observer) viewpoint occurs also in other attitudes; for episodic memories, he dismisses it with a spurious argument: “one cannot remember seeing oneself from a different perspective

¹⁴ McCarroll (2018, ch. 4) offers a detailed critical analysis of Vendler’s views with which the objections below mostly agree.

¹⁵ Williams (1973, p. 36) makes this point; cf. also Recanati (2007, p. 195) and McCarroll and Sutton (2017, p. 120).

simply because it is impossible to have seen oneself from an outside perspective” (Vendler, 1979, p. 169). Sutton (2010, Sect. 3) has a good discussion of Wollheim’s similar argument. Vendler’s argument makes two dubious assumptions. The first is the implicit presupposition that correct episodic memories shouldn’t just represent the same events as the perceptual experiences that are their causal source. Correct episodic memories, Vendler assumes, should *only* reproduce features of those perceptual experiences, even if they don’t preserve all its details in fully precise ways; they shouldn’t just be *truthful*, but also “authentic” (Michaelian & Sant’Anna, 2022, pp. 836–838). This is the “preservative” view of memory: perceptual experiences leave “traces” reproducing some of their contents, in some detail, retrieved later in recollection episodes (Michaelian & Sant’Anna, 2022, p. 844). As said above, I take this view to have been empirically refuted by now. Observer memories illustrate the reconstructive character of episodic memory, leaving nonetheless open the possibility that a (normative) causal condition distinguishes them from imaginings.

Second, McCarroll (2018) argues that Vendler’s argument doesn’t work also because the claim that “it is impossible to have seen oneself from an outside perspective” is wrong. The initiating experience might have included sufficient information to validate the observer memory. McCarroll makes a forceful case for this in his book, but Michaelian and Sant’Anna (2022) question this claim; following them, I’ll grant Vendler’s second assumption. But, as said, there are very good reasons to dismiss the first, including phenomenal evidence. I won’t go into the details here; Debus (2007), Sutton (2010), McCarroll and Sutton (2017), McCarroll (2018) and Sant’Anna (2018, p. 6) articulate compelling responses to dismissals of observer memories. The view is too revisionist; both the folk (Dranseika et al., 2021) and empirical research grant that observer experiences can be episodic memories.

In this section I have presented my main assumptions about personal memories, and the mnemonic experiences I take to constitute them. I extend them the methodological claims about IEM made in the previous section. I hypothesize that they identify a psychological natural kind. We are thus not free to stipulate their character; there is right or wrong about this, in particular about whether observer “memories” belong in the category, and hence also about how they are best understood. There of course also are better and worse theoretical accounts of their nature.

3 Episodic memory is not *logically* IEM

As said, Shoemaker (1970) argues that the memory-based judgment ‘I saw John yesterday’ is only *de facto* IEM. Evans (1982) challenged Shoemaker’s views. The debate was conducted based on weird science-fiction cases involving brain-splitting and brain-transplants.¹⁶ Pryor (1999, pp. 290–297) defends Shoemaker’s side against Evans’s arguments—for perception (*I see John*) also by appealing to more mundane situations involving mirrors and changes in point of view. He (1999, p. 304, fn. 54) also defends

¹⁶ The cases may not feel that bizarre. Folk viewers of *Blade Runner* (Scott, 1982) don’t manifest any difficulty in accepting that this admission by Rachael presents a coherent possibility: “I remember [music] lessons, but I don’t know if it is me or Tyrell’s niece” (Deckard has just told her that her private memories were “implants” from those of Tyrell’s niece).

Shoemaker's view for perception by using science-fiction examples—H. G. Wells' "Davidson's eyes" or *Avatar*-like cases, on which one is occasionally "hooked up" to someone's else body (O'Brien, 2007, pp. 38–39; Higginbotham, 2010, pp. 262–263). These conditions support the conjecture that someone may indeed see John, but it is my avatar, not me. If those cases establish that perception is not logically IEM, they also show that episodic memories fail to be so, on similar grounds; because the relevant memories might derive from perceptual experiences that come from the body that one was "hooked up" to at the time.

I'll summarize the gist of Evans's discussion because some of the considerations in debates on memory we'll confront are analogous. Remember the characterization we offered in Sect. 1 of EM and its possibility or otherwise, VEM versus IEM. To show that a source of evidence (perception, memory) at most warrants de facto IEM, we must describe a situation (which may only obtain in remote possible worlds, for the error need not occur in nearby worlds given that the judgment is de facto IEM) that defeats justification for the belief *a is F*. Related justification exists, but it comes from a source other than *a*: the judgment *a is F* is defeated, perhaps false, while its existential "part" (entailment, really) *that someone is F* remains justified.

Now, Pryor rightly insists that the defeater to be provided for *a is F* must be *undercutting* but not *additive*. An undercutting defeater for *p* undermines the justification for believing it; and *additive* defeater gives, in addition, positive evidence for *not-p* (Pryor, 1999, p. 284). If additive defeaters are allowed, there would not be *any* cases of IEM, not even *I am in pain* when based on introspection. Smith (2006, p. 279) argues for this skeptical result, relying on additive defeaters:

[...] suppose you experience what you take to be a pain. Now suppose someone reliably tells you that you are not really experiencing a pain but an itch, and that this has been caused by someone else's suffering a real pain. Whenever they experience a pain, they press a button which causes you to feel a pain-like itch.

This illustrates why additive defeaters should be disallowed. To show that a judgment is VEM it must be established that the justification for *a is F* *already in place before exposure to the defeater* still justifies the existential entailment, *someone is F*, after the defeat of *a is F*.¹⁷ Intuitively, it is not possible to undercut the introspective justification for *I am in pain* without defeating also *someone is in pain*. Smith's story only offers *new* justification for the existential claim, sneaking it with the defeater—it tells us that when we think we feel pain, someone else does instead. This doesn't show that the initial justification for *I am in pain* given by the subject's nociceptors provides justification for *someone is in pain* that can survive its defeat. Imagine that we are just told that data from a brain scan suggests that it is an itch that we feel. If this defeats the singular claim, it defeats with it its existential "part". We are left with

¹⁷ Hamilton puts it well: "IEM arises only if the *original* justification cannot be cited as support for "Someone is F" when the assertion is doubted for any reason at all" (Hamilton 2007, p. 411), my emphasis; cf. also McGlynn (2021, p. 2298). This is the *desideratum* that Palmira (2020, p. 3840) calls "Preservation of Grounds".

no good reason to think that the pain we thought we were feeling is in fact someone else's.¹⁸

Evans's main complaint about Shoemaker's argument for the merely de facto IEM character of memory is that his cases are similarly inapt. Evans (1982, pp. 144–145) argues that, if we get apt (non-additive) reasons to disbelieve our memories, it would be absurd to stick to the hypothesis that we were nonetheless experiencing the apparent memories of someone else. The proper conclusion should rather be that we are more prone to memory illusions than we thought. Only by smuggling into the defeater some additional information—some story to the effect that we are experiencing the true memories of someone else, like the one Deckard tells Rachael in *Blade Runner*, fn. 16—could we still sensibly endorse the existential “part” of the original claim. Pryor (1999, pp. 294–296) addresses Evans's points convincingly enough in my view. But more ordinary cases—like one involving mirrors Pryor (1999, p. 297) that Pryor devises to show the merely de facto character of the IEM of perception – would be helpful. I'll now argue that recent findings about the constructive nature of personal memory offer us precisely that regarding their IEM.¹⁹

Let me first introduce the distinction between “vicarious” and “disputed” “memories”. Vicarious “memories” (Pillemer et al., 2015) are not really memories. They are first-personal imaginings of specific episodes reported by others, whose subjects “are fully aware that the episode happened to someone else” (Pillemer et al., 2015, p. 234) and hence, I take it, don't experience them as memories.²⁰ This is in contrast with “disputed memories, where two individuals (oftentimes twins) both claim to be the protagonist in a single past event” (Pillemer et al., 2015). At least one of the disputed mnemonic experiences must be incorrect; while, as imaginative episodes that are experienced as such, vicarious “memories” aren't wrong just because they don't have their source in a perceptual experience of the subject. Let us consider now two stories in the literature that, I'll argue, depict consistent possibilities showing that memory is not logically IEM:

I am reminiscing with my colleagues about a philosophical meeting in which I made a brilliant objection to the speaker's thesis. However, none of us can remember any occasion on which I made that point. Then someone recalls Andrew

¹⁸ Remember Shoemaker's (1968, p. 563f.) “package deal” intuitively correct point, already quoted in Sect. 1: “[...] in being aware that one feels pain one is, tautologically, aware, not simply that the attribute *feel(s) pain* is instantiated, but that it is instantiated *in oneself*”.

¹⁹ Fernández (2021, pp. 646–647) articulates a train of thought to conclude that observer memories are not IEM, which I don't think pertinent because it depends on an idiosyncratic understanding of observer memories. On his view, what makes an episode an observer memory is not that the subject takes herself to be seen in the recalled event—the standard definition of observer memories, cf. the quotations from McCarroll and Sutton (2017) above and Nigro and Neisser (1983) below. It is rather that, *as a matter of fact*, one of the participants in the visualized event is the subject of the mnemonic experience, even if she doesn't identify herself as such. If (but only if) in fact it is the subject, it is an observer memory (Fernández 2021, pp. 645–646). On the standard characterization, as a referee put it to Fernández, “there is no need for self-identification in observer memory because, when one has an observer memory, there is no open question as to who is being visualized as being part of the relevant episode” (Fernández 2021, p. 646 fn. 15). On Fernández's view it is not a given that subjects identify themselves in observer memories, they just may do it; and this is how he envisages their liability to EM.

²⁰ The issue here is how memories (mnemonic experiences really, fn. 8) are identified by the subject, on which as I said I assume they are mostly reliable, cf. fn. 9.

Brennan making the very same point in a discussion which bears striking similarities to the one I have described. I am persuaded that this indeed was the discussion in question, and that I have misidentified the person who made the objection. “I made a brilliant objection” thus seems to be a mis-remembering—a personal memory-judgment with an incorrect detail, mistaken due to an error in identification. It seems natural to say that my reminiscence was based on information that was garbled, resulting in a misidentification of the maker of the objection as myself. Thus memory-judgments are not guaranteed IEM (Hamilton, 2007, pp. 413–414)

[I]magine that a subject S has the apparent memory of a soccer game and forms on the basis of it the first-personal judgment: ‘I then saw Giroud scoring’. But also imagine that it was a distinct subject S’ who actually attended the game and told S about it, and that because of this testimony S has ended up entertaining the merely apparent memory that he himself attended the game. S would then misattribute to himself the property of having attended the game, but he would not misidentify himself as S’, which is an error that seems to be excluded (Perrin & McCarroll, 2023, p. 301)

Let me speculate on how Hamilton’s mnemonic experience might have been based on “garbled” information, and how S’s derived from the testimony of S’. This is not meant to be a correct account of actual cases of the formation of erroneous memories, which would need empirical support. For my purposes I only need to claim that they articulate coherent *possibilities*. When critically addressing Evans’s arguments against the view that Shoemaker’s cases establish that memory is only de facto IEM, Pryor (1999, p. 290) confesses that he “doesn’t have much positive argument to present in support of the claim” that the subject’s defeated apparent memories still offer her knowledge of the existential claim; this “just seems intuitively plausible to me” (Pryor, 1999) We can do better. We can suppose that S’s erroneous memories in Perrin and McCarroll’s (2023) story derive from vicarious “memories” (imaginings really, as I said) that the subject (as they put it) ends up misidentifying as memories. The mechanism might be analogous in Hamilton’s case: a perceptual experience correctly representing Brennan raising the brilliant objection gets appropriated as a (self-complacent) memory of oneself asking it, perhaps an observer one.²¹

When the subjects in the stories get their protagonism in their memories disputed, they still may be confident about aspects of the information they purport to provide—say, true details about the room, the location of the person raising the objection that the subject wrongly identified as herself and other colleagues, the identity

²¹ James’s (2021, pp. 193–194) similar *Aunt Gayle* case can be understood analogously. Even though the relevant states are first imaginings, they end up being mnemonic experiences; for this is, by default, a matter of how their subjects identify them, cf. previous fn. The ontology of true memories might be relational, as the factivity of ‘memory’ suggests, cf. fn. 8. A real memory should perhaps derive through a non-deviant causal chain from a perceptual experience with a related content had by the same subject. But it is *mnemonic experiences* identified on more intrinsic grounds that a critical discussion of the IEM of memory must feature—which is why Shoemaker introduced his related technical notion of q(uasi)-memories). Shoemaker and Evans were not debating whether true memories as just envisaged are logically IEM; they trivially are. At issue is whether judgements based on mnemonic experiences are logically IEM—i.e., to put it in these terms, whether mnemonic experiences must be real memories.

of the speaker to whom the question was posed, the contents of her talk, her answer to the question, and so on. Facts such as the subject's phenomenally grounded confidence in such details, their reliability, safety, or their counterfactual dependence on what transpired at the meeting may suffice (given the right epistemology of memory, whether dogmatist, reliabilist, or whatever) to preserve the subject's justification for the existential claim: ok, I didn't ask the question, but someone did. These non-science fiction, rather ordinary cases show that, as Shoemaker thought, memories are only de facto IEM, given their reconstructive character that observer memories contribute to revealing.²²

The *metasemantic* account of IEM presented in García-Carpintero (2018) and Palmira (2020) underwrites these counterexamples to the logical IEM of memories. It may well be that, along the lines Peacocke (2014) suggests sketched above (Sect. 3), self-identification with a character in episodic memories—be they *field* or *observer*—is, in normal conditions, part of the reference-fixing information for the referent of 'I'. This would explain on the metasemantic account why these attitudes are de facto IEM. Still, the only logically non-negotiable aspect of that reference-fixing is the description provided by the token-reflexive self-reference rule. Hence, in non-standard conditions like those that our stories present, the subject might consistently retreat to the option that she can just be identified as the experiencer of the relevant mnemonic episode, while it is an open question whether she really had the properties ascribed to the individual represented in them as herself.

4 Arguments that episodic memory is IEM

I'll now critically examine recent arguments for the IEM of memory. Hamilton uses the sort of consideration by Evans outlined at the start of Sect. 3 to dismiss the ostensive conclusion of his story: without smuggling information into the defeater, making it *additive* instead of *undercutting*, "the idea that, if it was not myself, then it must have been someone else who made the objection, would just be a stab in the dark" (Hamilton, 2007, p. 414). But his argument for this claim overlooks the considerations for the reliability of the existential claim I offered two paragraphs back. He says that it "would not be justified if I initially thought that it was myself who had made the objection, because the detailed descriptions are in part third-personal, and I could not adopt the appropriate third-personal stance towards actions which are my own". This plainly ignores observer memories, for I do adopt in them an "appropriate third-personal

²² Morgan (2024) defends that episodic memories are IEM from two arguments, one that relies on Coliva's "background identity presupposition" view of VEM, and another on the notion of wh-misidentification. Although my account does assume such presuppositions, unlike Coliva's my discussion focusses on wh-misidentification, which I consider the fundamental phenomenon. As I also said (fn. 2), I think Morgan's is a good account of wh-VEM. The crucial point to resist his arguments against the wh-VEM of memory is the account that this paragraph offers for how the subject might be justified that someone asked the question, even though his (perhaps observer) mnemonic experience that it was he who did is proven inaccurate. Note by the way that if my cases are not just coherent counterfactual possibilities, but actual processes, this wouldn't establish that memory is not even de facto IEM. This depends on the proper way of characterizing that notion, and the epistemically relevant facts, see the next paragraph. It might be that in clear-cut situations constituting an epistemic default, memories are fully reliable.

stance” to actions I take to be my own. He also argues that the existential claim is not a credible one to make “since for it to be true, the subject’s memory-experience would have to change in order to accommodate new evidence—thus invalidating its status as memory-experience” (Hamilton, 2007). But if it was an observer memory, the only revision needed is the misidentification of the protagonist. Hamilton appears to assume the preservationist view that if any details in the attitude go beyond what the original perception represented, it cannot be a mnemonic experience.

Fernández (2014, 2021) has an argument for the IEM of memory. While Hamilton’s (2007, 410) pre-theoretical characterization of IEM is the one I am assuming, Fernández adopts one from Shoemaker that only specifies a subclass of the kind defined in Sect. 1, which he applies in ways I find inadequate. For any property P and mental state M, Fernández proposes, a judgment *I had P* made on the basis of mental state M is IEM relative to M just in case it is impossible that there is a subject S such that: (i) M represents S as P; (ii) M is fully accurate; (iii) I mistakenly think that I am identical to S, and (iv) my judgment that I had P is false as a result of (iii) (Fernández, 2021, p. 643). This characterization, close to Pryor’s (1999) *de re misidentification*, assumes that, in making judgments featuring errors of misidentification, one still has *de re* attitudes about an object truly instantiating the wrongly self-ascribed property. I agree with Pryor’s (1999, p. 293) reply to a related argument by Evans for the IEM of memory that this assumption is too strong. As I indicated in Sect. 1, in agreement with many other researchers, what is in general required for an EM is that the original evidence justifies the existential claim even if the judgment is wrong, whether an alternative singular judgment is also justified. However, this may obtain in the cases I am offering, and hence this discrepancy doesn’t affect our debate.

Fernández’s argument raises more relevant concerns. He (Fernández, 2014, 2017) defends a “factive” account of the content of episodic memories, on which “what we remember when we have a memory experience is a perception of a certain event; the objective event that we claim to remember in virtue of having that memory experience” (Fernández, 2014, p. 383). Then he argues for an “extrinsic” account of the content of the recollected perception, on which it represents an objective event as presented to the perspective of the perceiver, who is thus also part of the content: participants in the event are *in front* or *behind* the subject, at a certain distance, and so on. Of course, the experience targeted by an episodic memory is, on his view, only represented in it as a perception; no such perception might have occurred. But the sort of mistake required for EM, Fernández argues, is excluded, because conditions (i) and (ii) in his definition above cannot jointly obtain. For the memory to be “fully accurate” as (ii) requires, the represented perception must be one of its subject; (i) would then fail to obtain: “if my judgment happens to be false because the person who actually [had P] turns out to be someone else, then we should no longer consider my memory to be accurate” (Fernández, 2014, p. 388; 2017, p. 658).

Even if we grant Fernández’s view that all memories represent a previous purported perception of an event as presented from the (“field”) perspective of the recollecting subject (which we shouldn’t), the argument outlined is too quick. Even though he doesn’t adopt a fully preservationist view, Fernández can only dismiss the *Avatar* or *Davidson’s eyes* cases mentioned above, as his argument does, by interpreting the “full accuracy” condition (ii) in his definition of IEM in an unduly demanding way. On my

characterization of IEM what matters is whether the existential claim we obtain by abstracting away the subject is still justified by the mnemonic evidence, which, as Pryor (1999) argues, might be the case in *Avatar* cases.²³

We can however also put this aside for our purposes, because the most salient difficulty of Fernández's arguments from my viewpoint here is that the account of the content of memories he develops feels just like a sophisticated validation of Vendler's, Wollheim's or Hamilton's assumptions, which wrongly dismiss observer memories. Readers of Fernández (2021) might wonder why the paper devotes a good number of (interesting, to be sure) pages to investigating whether observer memories make Shoemaker's case that memories are not IEM. Shouldn't he just stick to his account of memory, on which all genuine instances are field memories? And then he would be done, because he takes himself to have also established that all such memories are IEM, with the argument outlined two paragraphs back.²⁴ Indeed, his account entails that when we recall *any* event, we *always* "see the scene as we originally saw it from a first-person or 'field' perspective" (McCarroll & Sutton, 2017, p. 114).

Although this objection is ultimately well-taken (cf. fn. 29 below), so stated it overlooks an aspect of observer memories we haven't yet explicitly discussed. We visualize ourselves in them an "external" participant in the event; hence, there is a perspective from which the visualization occurs (Lin & Dranseika, 2021, p. 406; Sutton, 2010, p. 28). McCarroll and Sutton say in the quotation above that it presents the past event "from a position we didn't occupy at the time of the original episode" (McCarroll & Sutton, 2017). This would of course be factually true if the observer mnemonic experience is correct: the subject was not in fact occupying the origin-of-perspective position. But it might be incorrect, and in fact the subject might have occupied it; this is how I imagined the Hamilton-inspired counterexample to the IEM of memory in the previous section. It shouldn't be in any case a necessary condition on observer memories that the subject is not presented in them as occupying also the origin-of-perspective position.²⁵ In fact, Nigro and Neisser make the opposite assumption in the paper that triggered contemporary research on the topic:

In some memories one seems to have the position of an onlooker or observer, looking at the situation from an external vantage point and seeing oneself "from the outside." In other memories the scene appears from one's own position; one seems to have roughly the field of view that was available in the original situation and one does not "see oneself" (Nigro & Neisser, 1983, 467-468)

Nigro and Neisser assume in this characterization that the perspective in observer memories is also that of the observed subject. Observer memories must have the

²³ We can (and should) interpret the definition of IEM by Shoemaker that Fernández adopts as allowing for this. Even if the subject figures in the content of the episodic memory as Fernández claims, the existential "part" of that content might be "fully correctly" instantiated by its witness (my avatar, say), whom I am wrongly taking to be myself; hence on this reading both (i) and (ii) in the Shoemakerian definition of IEM that Fernández borrows might be satisfied.

²⁴ Michaelian (2021, p. 9536) and Perrin and McCarroll (2023, p. 307) express related concerns; see also fn. 29 and 30 below.

²⁵ McCarroll (2018) argues that the geometrical origin of perspective in observer memories is left unoccupied. Even if a possible option, I don't think that this is necessary either.

paradoxical character of “out of body” experiences, in which one represents oneself as occupying two locations at once. Lin and Dranseika (2021) report experimental data in which a large proportion of subjects describe their observer imaginings in that way; but not all do. I think that the more accurate description consistent with the reports is that observer memories present the subject as one of the participants in the event visualized from a given perspective, leaving it open whether it is occupied (and, if so, by the subject); it may just be that of a device like a camera.²⁶ In field memories in contrast the subject always locates himself only at the origin of the visualizing perspective—the one he should have had at the originating perceptual experience. Fernández may think that, although his theory always places the subject at the location she has in field memories, some would count as observer memories on this more neutral characterization.

Now, although I agree with him that we shouldn’t assume that the origin of perspective in an observer memory is *not* the subject’s, we should equally reject the view that it *must be* the subject’s. This is highly relevant for the appraisal of his argument that observer memories don’t help to make Shoemaker’s case against the (logical) IEM of memory. He (Fernández, 2021, pp. 648–652) considers cases like Hamilton’s meeting that I invoked in Sect. 3 to this end: I visualize myself making the brilliant objection; but after being challenged I allow the possibility that it was Brennan who made it. Now, how about the perspective from which the situation is visualized, was it in fact occupied, and, if so, was it me who occupied it? For all I think we are entitled to assume, the perspective might just be that of a camera; the case that memory is not IEM has already been made if the situation is possible, the identity of the visualizer (if any) being irrelevant. On Fernández view, however, I must also represent myself in the episode as the holder of that perspective.²⁷ If I wasn’t in fact occupying it, the memory is also wrong on that count; but then, according to Fernández, it fails to establish that memories are not IEM, because his condition (ii) is falsified: the representation is not “fully accurate” (Fernández, 2017, pp. 648–650).

This is just the previous argument for the IEM of memory I already dismissed. Observer memories make the dismissal more compelling because they need not be inaccurate for this reason. Observer memories don’t necessarily present the subject as doing the visualizing in the scene; as pointed out above, the relevant perspective might be unoccupied. Against this, Fernández provides a question-begging argument:

[...] why is that side of the remembered passenger’s body the side that I visualise in virtue of having my observer memory? The answer to this question seems to be that the left side of their body is the side that I would have been able to see, had I been sitting behind the front passenger’s seat, given their position in the car relative to mine. It seems, therefore, that my observer memory represents not only the position of the remembered passenger in the car [...], but also my own position in the car. (Fernández, 2021, p. 648)

²⁶ Cf. Perrin and McCarroll (2023, p. 306) for a characterization congenial with this view.

²⁷ This is what Michaelian (2021, p. 9534) calls in his discussion of Fernández’s argument the “content” assumption. He (Michaelian, 2021, pp. 9538–9539) offers compelling reasons that mine parallel to reject it, concluding also that observer memories show that memory is not IEM; see also fn. 29.

But even if the counterfactual (“the side that I would have been able to see ...”) is true, the conclusion that the observer memory represents not only the position of the (observed) passenger I identify as myself, but also my own (visualizing) position doesn’t follow; as said, the visualizing might be ascribed to someone else, or to a camera.²⁸

In any case, Fernández should grant that I might in fact have been the one visualizing Brennan making the brilliant objection from the remembered perspective. This is the situation that I assumed when I presented the case in Sect. 3. On this understanding, the episode cannot be incorrect by Fernández’s lights. Shouldn’t he then grant that this possibility does show after all that memory is not IEM? In response, he questions that, now understood in this way, the episode is a memory at all. It is not an observer memory, he claims (Fernández, 2021, p. 650), relying on his idiosyncratic characterization, cf. fn. 19: if the subject is in fact the one doing the visualizing, he cannot have been any of the participants visualized in the represented events; hence, on his definition, it is not an observer memory. Crucially, he also argues that, depending on the correct metaphysics of memories, it may not even be one. He advances a plausible principle, which may exclude the mental episode we are considering from being a memory:

One might argue that one of the conditions that a mental image representing some scene must satisfy for it to qualify as a memory of the scene is that the mental image must dispose us to believe that the scene took place in the past. This is, after all, one of the ways in which, typically, memories are different from episodes of imagination (Fernández, 2021, p. 651)

I find this “belief” assumption—as Michaelian (2021, p. 9532) calls it—plausible; but it creates havoc when combined with Fernández’s view that the visualizer in all observer memories is the subject. The episode we are considering would thus represent the subject as being also the one doing the visualizing—the one visualizing the person asking the brilliant question. But of course, before being challenged by her colleagues, the subject doesn’t believe this. This being an observer memory on the standard definition (albeit not on Fernández’s), she takes herself to be the one who is seen asking the brilliant question; and, being rational, she doesn’t take herself to have been in two different locations at the same time. The “belief” principle is thus violated with respect to an aspect of what Fernández’s account takes to be the content of the episode and, if it is indeed a necessary condition for memories, the episode wouldn’t be one.²⁹

²⁸ In personal communication, Fernández tells me that he meant to understand the “I” who occupies the perspective from which the visualizing occurs in minimal terms, deprived of many of my bodily features, beliefs and so on – simply as a visual perspective that determines relations like “to the left”, “up”, and so on. I doubt that this is consistent with Fernández’s application of the view that the perspective is the subject’s to raise doubts that it is a memory at all in the argument I am about to discuss. In any case, as said I find it fully compatible with my intuitions that the perspective is just that of an impersonal recording device.

²⁹ As Michaelian (2021, p. 9536) points out, in fact Fernández’s *belief* and *content* assumptions together entail that *no* observer memories, as standardly understood, are memories: not just when the remembering subject was the one doing the visualizing in the recalled episode, but also when the recollection was correct and she was the one asking the brilliant question, and when she was neither one nor the other. In observer memories as standardly understood the subject might not believe that he is the one doing the visualizing, whether or not he actually was in the recalled event; but on Fernández’s account he should do, because that

I understand that Fernández takes this at least to pose a challenge for the Shoemaker line. From my perspective, the challenge immediately crumbles; to repeat, this is just a sophisticated articulation of a view like Vendler's, which we have very good reasons to reject. Observer memories need not present their subject as occupying both the "external" position in the represented event and the visualizing perspective. Their subjects might thus be rationally prepared to fully believe their content—that they, presented in the scene as visualized by someone or something, asked the brilliant question—and hence they count as memories even if we grant Fernández's "belief" principle. Elaborated as I suggested, cases like those Hamilton and Perrin & McCarroll describe establish that genuine memories are at most *de facto* IEM.

For all we can say at this stage in our research, episodic memories, observer memories, and IEM are real conditions for which we have well-supported characterizations that capture central cases of the phenomena, and theoretical accounts that are to be appraised relative to how they account for them. Fernández's arguments assume characterizations of observer memories and IEM that we are not free to stipulate; the former is idiosyncratic, and both are problematic. The characterizations I and many others have been assuming appear to capture better the shape of the phenomena.³⁰ In addition, Fernández relies on a theoretical account of memories validating strong "authenticity" requirements that we don't have good reasons to accept, among other reasons because observer memories as standardly understood question it.

I will conclude by briefly examining two other recent arguments that episodic memories are IEM; I'll argue that they also rely on questionable accounts of IEM. Lin (2020, p. 410) declares: "The issue of whether memory-based judgments are IEM concerns whether there is room for judgments that are justified by our memory to be wrong in a particular way: can we be wrong about the subject of such a judgment (i.e., the first-person pronoun)?" She then goes on to assume a continuist view of episodic memories, on which they are essentially imaginings. On this basis, she argues that both field and observer memories are IEM:

Experiencing the protagonist as oneself admits no doubt because there is no simulated subject prior to the identification. The remembered protagonist is created simultaneously as the identification is established. The implication is that judgments based on episodic memory are IEM. The immunity, however, stems from the self-simulating nature of identification, rather than the inheritance of the extrinsic property from perception as the inheritance view suggests" (Lin, 2020, p. 419)

The latter claim correctly rejects Fernández's (2021) argument for the view they share, which she takes to be based on a *inheritance* (preservationist, in my terms) view of memory; she favors instead a *reconstructivist* one. Now, although I embrace

Footnote 29 continued

he did is necessarily an aspect of the content of the memory. Thus, like Vendler and Wollheim even if on more elaborated grounds, Fernández might just declare that observer memories as standardly understood are not memories on his views about their contents, if the *belief* principle is assumed.

³⁰ On similar grounds, James (2021, Sect. 4) argues instead that Fernández's conclusion that memories are IEM is uninteresting, because it crucially depends on his definitions of *memory* and *IEM*. Given the methodological stance I advanced in Sects. 1, 2, I think that my appraisal is more accurate; no theorist is in a position to simply stipulate how these notions are to be understood.

a form of discontinuism, I take it to be compatible with reconstructivist views. But this is irrelevant for what I take to be the obvious problem with Lin's argument. I have already granted that self-identification in an observer memory (as much as in a field one, or in fact in any form of self-reference) is not a matter of identifying oneself with a previously given object, for which one has an alternative description. She is right that "[n]o mental action is required to identify with oneself, as it does not require the subject to do something to establish the identification with any pre-existing simulated subject" (*ibid*). But this doesn't establish the IEM of episodic memories; it is instead a different general feature of self-reference with 'I', obtaining both in IEM and in VEM cases.

The feature of 'I' that Lin wrongly takes to establish IEM is one pointed out by Anscombe that Coliva (2003) calls "the Real Guarantee".³¹ Other referential devices allow for a split between *speaker's* and *semantic* referent. I utter 'this is a picture of one of the greatest twentieth century philosopher' pointing behind me to the place where until recently there was a picture of Carnap which, unbeknownst to me has been replaced with one of Agnew (Kaplan, 1970). In one sense (*speaker's reference*), I refer to the Carnap picture I mean to refer to; in another (*semantic reference*), I refer to the one now behind me. Consider in contrast the case in which I see a picture of what I take to be myself looking at John and say on that basis 'I saw John'. If it was in fact someone else, there is *no sense* in which we would say that I have referred to that person with 'I'. As Anscombe puts it, "'I...' guarantees not only the existence but the presence of its referent. It guarantees the existence because it guarantees the presence".

Lin is thus right that to achieve reference with a use of 'I' "[n]o mental action is required to identify with oneself ...". The referent is guaranteed to be the "present" one, i.e., the one deploying the 'I' in the act. This is why we cannot "be wrong about the subject of such a judgment (i.e., the first-person pronoun)". But adopting this as a characterization of IEM is misguided, because it makes all uses of 'I' IEM—including paradigm cases of VEM judgments, as in the illustration I have been providing, 'I saw John' based on self-identification in a picture. I am referring to myself there and not to the person in the picture, but the claim is still EM.³²

Perrin and McCarroll (2023) argue for a complex account on which some memories are IEM, but some others are not. I agree with the VEM cases they provide, and their considerations about them (Perrin & McCarroll, 2023, pp. 318–319); we are thus in agreement on the main issue at stake here. I lack the space to go into the additional complexity of their view. I'll just note that in my view it is compromised because they deploy a characterization of IEM analogous to Lin's. They say that, when they are VEM, "the content of episodic memories is identity neutral and the judgment the remembering subject forms includes an identification component, which is possibly mistaken" (Perrin & McCarroll, 2023, p. 301). Thus, when at the end of the quotation I provided in Sect. 3 with the Giroud example, they say "S would then misattribute to himself the property of having attended the game, but he would not misidentify himself

³¹ Christofidou (1995) defended a characterization of IEM like Lin's.

³² See Coliva (2003) for a compelling critical discussion along these lines of previous instances of the mistake that I am ascribing to Lin, including Christofidou's.

as S', which is an error that seems to be excluded", they take this to establish that this memory is IEM, against what I argued. In this argument they seem to be understanding the notion as Lin does: VEM requires a "neutral", independent identification of the relevant individual; cases in which subjects wouldn't identify themselves as being other than the referent of the instance of 'I' they utter would have guaranteed IEM. This is an invalid line of reasoning, because it would discard as VEM paradigm instances thereof.

In this paper I have offered an argument that the reconstructive character of memories, which observer memories illustrate, allows more mundane possibilities to make Shoemaker's case that memory is only de facto IEM than previous debates featured. I have relied on an elaboration of Shoemaker's characterization of IEM, and on what I take to be the best available theoretical account of the phenomenon. I have critically examined recent arguments for Evans's view that memory is logically IEM to a good extent on methodological grounds, by questioning assumptions they make about how to characterize memory, IEM, or both.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. Financial support for my work was provided by MICIU/AEI/<https://doi.org/10.13039/501100011033>, research projects PID2020-119588 GB-I00, CEX2021-001169-M, and through the award "ICREA Academia" for excellence in research, 2018, funded by the Generalitat de Catalunya. The paper was presented at the Grenoble "Reference in Remembering" workshop, June 2022, and received helpful comments there. Thanks to Jordi Fernández, Chris McCarroll, Kourken Michaelian, James Openshaw, Denis Perrin, André Sant'Anna, and the referees for this journal for very useful comments and suggestions, and to Michael Maudsley for the grammar revision.

Declarations

Competing interests The author of this paper doesn't have any conflict of interest regarding the reported research.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Anscombe, G. E. M. (1975). The first person. In S. D. Guttenplan (Ed.), *Mind and language* (pp. 45–65). Clarendon Press.
- Campbell, J. (1997). Sense, reference and selective attention. *Proceedings of the Aristotelian Society*, 71, 55–98.
- Christofidou, A. (1995). First person: The demand for identification-free self-reference. *The Journal of Philosophy*, 92(4), 223–234.
- Coliva, A. (2003). The first person: Error through misidentification, the split between speaker's and semantic reference, and the real guarantee. *Journal of Philosophy*, 100, 416–431.

- Coliva, A., & Palmira, M. (2024). Immunity to error through misidentification: Some Trends. *Philosophical Psychology*.
- D'Ambrosio, J., & Stoljar, D. (2021). Vendler's puzzle about imagination. *Synthese*, 199(5–6), 12923–12944.
- Debus, D. (2007). Perspectives on the past: A study of the spatial perspectival characteristics of recollective memories. *Mind & Language*, 22(2), 173–206.
- Dings, R., McCarroll, C. J., & Newen, A. (2023). Situated authenticity in episodic memory. *Synthese*. <https://doi.org/10.1007/s11229-023-04309-w>
- Dokic, J. (2014). Feeling the past: A two-tiered account of episodic memory. *Review of Philosophy and Psychology*, 5(3), 413–426.
- Dranseika, V., McCarroll, C. J., & Michaelian, K. (2021). Are observer memories (accurate) memories? Insights from experimental philosophy. *Consciousness and Cognition*. <https://doi.org/10.1016/j.concog.2021.103240>
- Evans, G. (1982). *The varieties of reference*. Clarendon Press.
- Fernández, J. (2014). Memory and immunity to error through misidentification. *Review of Philosophy and Psychology*, 5, 373–390.
- Fernández, J. (2017). The intentional objects of memory. In S. Bernecker & K. Michaelian (Eds.), *The Routledge handbook of memory* (pp. 88–99). Routledge.
- Fernández, J. (2021). Observer memory and immunity to error through misidentification. *Synthese*, 198, 641–660.
- Fernández, J. (2024). Memory and self-reference. *International Journal of Philosophical Studies*. <https://doi.org/10.1080/09672559.2024.2334461>
- Fine, K. (2007). *Semantic relationism*. Blackwell.
- García-Carpintero, M. (2015). De se thought. In S. Golberg (Ed.), *Oxford handbooks online*. Routledge.
- García-Carpintero, M. (2016). Token-reflexive presuppositions and the de se. In M. García-Carpintero & S. Torre (Eds.), *About oneself* (pp. 179–199). Oxford University Press.
- García-Carpintero, M. (2017). The philosophical significance of the *de se*. *Inquiry*, 60(3), 253–276.
- García-Carpintero, M. (2018). *De se* thoughts and immunity to error through misidentification. *Synthese*, 195, 3311–3333.
- García-Carpintero, M. (2024). Is conscious thought immune to error through misidentification? *Philosophical Psychology*. <https://doi.org/10.1080/09515089.2024.2351535>
- Gregory, D. (2016). Inner speech, imagined speech, and auditory verbal hallucinations. *Review of Philosophy and Psychology*, 7, 653–673.
- Gregory, D. (2023). Imagining a way out of dream skepticism. *Erkenntnis*. <https://doi.org/10.1007/s10670-023-00662-x>
- Hamilton, A. (2007). Memory and self-consciousness: Immunity to error through misidentification. *Synthese*, 171, 409–417.
- Higginbotham, J. (2010). On words and thoughts about oneself. In I. Stojanovich, F. Recanati, & N. Villanueva. (Eds.), *Context-dependence, perspective, and relativity* (pp. 253–282). Berlin: Mouton de Gruyter.
- Hunt, M. W. (2023). Do imaginings have a goal? *Global Philosophy*, 33, 11. <https://doi.org/10.1007/s10516-023-09658-2>
- James, S. (2021). Immunity to error through misidentification and the functionalist, self-reflexive account of episodic memory. *Estudios De Filosofía*, 64, 189–200.
- Kaplan, D. (1970). Dthat. In P. Cole (Ed.), *Syntax and semantics 9* (pp. 221–243). Academic Press.
- Kompa, N. (2023). Inner speech and 'pure' thought—do we think in language? *Review of Philosophy and Psychology*. <https://doi.org/10.1007/s13164-023-00678-w>
- Kratzer, A. (2021). Situations in natural language semantics. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy (winter 2021 edition)*. Stanford University.
- Lin, Y. T. (2020). The experience of being oneself in memory: Exploring sense of identity via observer memory. *Review of Philosophy and Psychology*, 11, 405–422.
- Lin, Y. T., & Dranseika, V. (2021). The variety and limits of self-experience and identification in imagination. *Synthese*, 199, 9897–9926.
- Martin, M. G. F. (2001). Out of the past: Episodic recall as retained acquaintance. In C. Hoerl & T. McCormack (Eds.), *Time and memory: Issues in philosophy and psychology* (pp. 257–284). Oxford University Press.

- McCarroll, C. J. (2018). *Remembering from the outside: Personal memory and the perspectival mind*. Oxford University Press.
- McCarroll, C. J., & Sutton, J. (2017). Memory and perspective. In S. Bernecker & K. Michaelian (Eds.), *The Routledge handbook of memory* (pp. 113–126). Routledge.
- McGlynn, A. (2021). Immunity to wh-misidentification. *Synthese*, 199, 2293–2313.
- Mehta, N. (2023). The common kind theory and the concept of perceptual experience. *Erkenntnis*, 88, 2847–2865.
- Michaelian, K. (2021). Observer memory and immunity to error through misidentification. *Synthese*, 198, 9525–9543.
- Michaelian, K., Perrin, D., & Sant'Anna, A. (2020). Continuities and discontinuities between imagination and memory: The view from philosophy. In A. Abraham (Ed.), *The Cambridge handbook of the imagination* (pp. 293–310). Cambridge University Press.
- 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.
- Michaelian, K., & Sant'Anna, A. (2022). From authenticism to alethism: Against McCarroll on observer memory. *Phenomenology and the Cognitive Sciences*, 21, 835–856.
- Morgan, D. (2019). Thinking about the body as subject. *Canadian Journal of Philosophy*, 49, 435–457.
- Morgan, D. (2024). Memory and identity. *Philosophical Psychology*, 2024, 1.
- Nigro, G., & Neisser, U. (1983). Point of view in personal memories. *Cognitive Psychology*, 15(4), 467–482.
- O'Brien, L. (2007). *Self-knowing agents*. Oxford University Press.
- Palmira, M. (2020). Immunity, thought insertion, and the first-person concept. *Philosophical Studies*, 177, 3833–3860.
- Palmira, M. (2022). Questions of reference and the reflexivity of first-person thought. *Journal of Philosophy*, 119(11), 628–640.
- Peacocke, C. (2008). *Truly understood*. Oxford University Press.
- Peacocke, C. (2014). *The mirror of the world*. Oxford University Press.
- Perrin, D., & McCarroll, C. J. (2023). Immunity to error through misidentification in observer memories: A moderate separatist account. *Philosophy and Phenomenological Research*, 106, 299–323.
- Perrin, D., & Sant'Anna, A. (2022). Episodic memory and the feeling of pastness: From intentionalism to metacognition. *Synthese*, 200, 109. <https://doi.org/10.1007/s11229-022-03567-4>
- Pillemer, D. B., et al. (2015). Vicarious memories. *Consciousness and Cognition*, 36, 233–245. <https://doi.org/10.1016/j.concog.2015.06.010>
- Prosser, S. (2012). Sources of immunity to error through misidentification. In S. Prosser & F. Recanati (Eds.), *Immunity to error through misidentification: New essays* (pp. 158–179). Cambridge University Press.
- Pryor, J. (1999). Immunity to error through misidentification. *Philosophical Topics*, 26, 271–304.
- Recanati, F. (2007). *Perspectival thought*. Oxford University Press.
- Recanati, F. (2009). *De re and de se*. *Dialectica*, 63, 249–269.
- Recanati, F. (2012). Immunity to error through misidentification: What it is and where it comes from. In S. Prosser & F. Recanati (Eds.), *Immunity to error through misidentification: New essays* (pp. 180–201). Cambridge University Press.
- Rice, H. J. (2010). Seeing where we're at: A review of visual perspective and memory retrieval. In J. H. Mace (Ed.), *The act of remembering: Toward an understanding of how we recall the past* (pp. 228–258). Wiley-Blackwell.
- Robins, S. (2020). Defending discontinuism, naturally. *Review of Philosophy and Psychology*, 11, 469–486.
- Rowlands, M. (2018). The remembered: Understanding the content of episodic memory. In K. Michaelian, D. Debus, & D. Perrin (Eds.), *New directions in the philosophy of memory* (pp. 279–293). Routledge.
- Sant'Anna, A. (2018). Episodic memory as a propositional attitude: A critical perspective. *Frontiers in Psychology*, 9, 1220.
- Seeger, M. (2015). Immunity and self-awareness. *Philosophers' Imprint*, 15(17), 10.
- Senor, T. D. (2022). The epistemology of episodic memory. In A. Sant'Anna, C. J. McCarroll, & K. Michaelian (Eds.), *Current controversies in philosophy of memory* (pp. 227–243). Routledge.
- Shoemaker, S. (1968). Self-knowledge and self-awareness. *Journal of Philosophy*, 65, 555–567.
- Shoemaker, S. (1970). Persons and their pasts. *American Philosophical Quarterly*, 7, 269–285.
- Smith, J. (2006). Which immunity to error? *Philosophical Studies*, 130, 273–283.

- Soteriou, M. (2020). The disjunctive theory of perception. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy (summer 2020 edition)*. Stanford University.
- Sutton, J. (2010). Observer perspective and acentred memory: some puzzles about point of view in personal memory. *Philosophical Studies*, 148(1), 27–37.
- Teroni, F. (2024). Memory identification and its failures. *Philosophy and the Mind Sciences*, 5, 28.
- Vendler, Z. (1979). Vicarious experience. *Revue De Métaphysique Et De Morale*, 84(2), 161–173.
- Verdejo, V. (2021). Perspectives on de se immunity. *Synthese*, 198, 10089–10107.
- Williams, B. (1973). Imagination and the self. *Problems of the self* (pp. 26–45). Cambridge University Press.
- Wittgenstein, L. (1958). *Philosophical Investigations*. G.E.M. Anscombe (tr.). Oxford: Basil Blackwell.
- Wright, C. (1998). Self-knowledge: The Wittgensteinian legacy. In C. Wright, B. C. Smith, & C. McDonald (Eds.), *Knowing our own minds* (pp. 13–45). Oxford Clarendon Press.
- Wright, C. (2012). Reflections on François Recanati's 'immunity to error through misidentification: What it is and where it comes from.' In S. Prosser & F. Recanati (Eds.), *Immunity to error through misidentification: New essays* (pp. 247–280). Cambridge University Press.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.