Immunity to error through misidentification and the functionalist, self-reflexive account of episodic memory*

Steven James
West Chester University of Pennsylvania, Pennsylvania, United States
E-mail: sjames@wcupa.edu

Recibido: 31 de enero de 2021 | Aceptado: 9 de abril de 2021
https://doi.org/10.17533/udea.ef.n64a10

Abstract: Fernández (2019) offers an account of the nature of episodic memory that marries two core ideas: (i) role-functionalism about episodic memory, and (ii) self-reflexive mnemonic content. One payoff of this view is that episodic memory judgments are immune to error through misidentification. Fernández takes this to reveal something important about the nature of one’s self-awareness in memory and our first-person conception of ourselves. However, once one sees why such judgments are immune in this way, according to the proposed account, the fact that they are immune becomes moot. While technically immune to error through misidentification, episodic memory judgments are not grounded in a way such that they have any interesting epistemological import for the subject (in contrast to other paradigms of such judgments), and any insights about our self-awareness and self-conception are directly derivable from the metaphysics of memory content alone.

Keywords: immunity to error through misidentification, memory judgments, episodic memory, self-awareness, metaphysics of memory

* This paper began as a talk at a workshop on Fernández’s book, Memory: a self-referential account, held at the Centre for Philosophy of Memory at Université Genoble Alpes. This work has been jointly funded by the Centre for Philosophy of Memory, and West Chester University of Pennsylvania

Cómo citar este artículo
James, S. (2021). Immunity to error through misidentification and the functionalist, self-reflexive account of episodic memory. Estudios de Filosofía, 64, 189-200. https://doi.org/10.17533/udea.ef.n64a10
Inmunidad al error de identificación errónea y la teoría funcionalista y autorreflexiva de la memoria episódica

**Resumen:** Fernández (2019) ofrece una teoría de la naturaleza de la memoria episódica que reúne dos ideas centrales: (i) el funcionalismo sobre la memoria episódica y (ii) el contenido mnemónico autorreflexivo. Una ventaja de esta teoría es que los juicios de memoria episódica son inmunes al error por identificación errónea. Según Fernández, esto revela algo importante sobre la naturaleza de la conciencia de sí mismo en la memoria y nuestra concepción en primera persona de nosotros mismos. Sin embargo, una vez que uno ve por qué tales juicios son inmunes, según el relato propuesto, el hecho de que sean inmunes se vuelve discutible. Si bien técnicamente son inmunes al error a través de la identificación errónea, los juicios de memoria episódica no se basan en una forma tal que tengan una importancia epistemológica interesante para el sujeto (en contraste con otros paradigmas de tales juicios), ya que cualquier conocimiento sobre nuestra autoconciencia y autoconcepción se derivan directa y únicamente de la metafísica del contenido de la memoria.

**Palabras clave:** inmunidad al error de identificación errónea, juicios de memoria, memoria episódica, auto-conciencia, metafísica de la memoria

**Steven James**

Is associate professor at the West Chester University of Pennsylvania. His specializations are analytic epistemology, philosophy of mind, cognitive science, and philosophy of language. He is currently completing a book-length project on the nature of remembering things.

**ORCID:** 0000-0002-8370-8746
Fernández (2019) offers a novel, interesting, and quite clever account of the nature of episodic memory that marries two core ideas: (i) role-functionalism about episodic memory—a mental state counts as a memory in virtue of playing a particular functional role in the cognitive system of the subject (chapter 2), and (ii) self-reflexive mnemonic content— the contents of episodic memories are quite ‘thick’ in that they are necessarily constituted by representations of all of the following: (a) the memory itself, (b) a past (objective) fact, (c) one of the remembering subject’s past perceptual experiences, and (d) a causal relationship between that past experience and the present memory (chapter 3). One payoff of this view is that judgments on the basis of genuine episodic memories are, despite appearances, necessarily free of a certain kind of error—those involving the misidentification of the self in memory. In other words, they are immune to error through misidentification, hereafter ‘IEM’ (chapter 6). Fernández takes this surprising result to reveal something important about one’s self-awareness in memory (p. 143) and our first-person conception of ourselves (p. 169). However, once one sees why such judgments are IEM, according to the proposed account, the fact that they are IEM becomes considerably less interesting. Episodic memory, while technically IEM, does not meaningfully ground judgments that have any special epistemological import for the subject (in contrast to other paradigms of IEM judgments). And any insights about self-awareness and self-conception are directly derivable from the metaphysics of memory content alone.

The article is structured as follows. In section 1, I introduce the notion of IEM, and review some reasons that states with such status could be theoretically important. In section 2, I present a case that common sense, and Fernández’s own functionalist account of episodic memory, would suggest is an instance of episodic memory that fails to be IEM in the relevant way. I then explain why it would not be surprising for episodic memory to fail to be IEM. In section 3, I explain why Fernández is, nevertheless, in a position to reject this conclusion by appealing to his account of episodic memory content in conjunction with the particular formulation of the IEM thesis he has adopted. In section 4, I show that this way of establishing that episodic memory is IEM strips the claim of its theoretical interest.

1. Immunity to error through misidentification (IEM)

Despite the well-known fallibility of human cognition, many philosophers are tempted by the thought that the scope of our ability to be mistaken is somehow limited; there are certain domains in which one’s judgments have some special kind of epistemic status. Of particular interest here is the claim that there are certain kinds of judgments which avoid the possibility of ‘mistaken identity’. In other words, there are judgments in which it is impossible to form, ‘a false belief because of a misidentification of the person or object about whom one made the judgment’ (Prosser and Recanati, 2012, p. x).
In the literature, such judgments are labeled immune to error through misidentification (IEM) and contrasted with judgments which are subject to error through misidentification. The least controversial candidates are those involving first-person judgments about oneself on the basis of occurrent conscious experiences, e.g. introspective judgments about one’s current conscious mental states, proprioceptive judgments. Borrowing again from Prosser and Recanati (2012, p. ix), two exemplars are, “‘I have a headache’ (where the judgment is made on the grounds that one feels one’s head aching) and ‘my legs are crossed’ (where the judgment is made on the basis of proprioception). In the first case, it is said, I could not be mistaken about whose headache I was aware of; and in the second kind of case I could not be mistaken about whose legs I felt to be crossed.” In contrast, judgments like, ‘Bill was rude at the party’ (on the basis of observing his behavior) are subject to error through misidentification because it is possible for one to mistake someone else for Bill.

Following Prosser (2012, p. 161), it seems that what makes a judgment susceptible to error is a certain kind of structure: (i) some object (a) is the causal source of information that leads to (motivates/justifies, etc.) the application of a predicate (F). (ii) The predicate (F) gets applied to an object (b). And error arises when a ≠ b. So, one should expect a formula for identifying judgments that are immune to such errors. Find those judgments, should there be any, in which the causal source that leads to the application of the predicate, and the object to which the predicate gets applied cannot fail to be identical.¹

Finding such types of judgments could be interesting for a variety of reasons. It might inform our understanding of the structure of justification —IEM judgments may have a kind of epistemological priority (Evans, 1982, pp. 181-182). It might inform our understanding of the semantics of indexicals and demonstratives (Evans, 1982; Recanati, 2007). It might inform our understanding of the structure of self-knowledge (Evans, 1982; Ismael, 2012; Merlo, 2017). And most importantly for present purposes, it might reveal something important about our awareness of ourselves in memory, and our conception of the self (Evans, 1982; Fernández, 2019).² As Fernández sees it, ‘[i]f memory judgments are IEM, then (…) in memory, one is aware of the subject who is remembered to have instantiated such-and-such properties as being oneself’ (2019, p. 143). More precisely, the IEM status of memory judgments would reveal, the fact that, in memory, we are presented to ourselves (…) as the bearers of extrinsic properties which were perceived in the past; properties such as occupying a certain spatial position or having a particular size relative to that of another object (…) [with a final lesson being] that our first-person conception of

¹ Cappelen & Dever (2013, pp. 130-33) make a reasonably compelling case that there is no philosophically interesting phenomenon of immunity to error through misidentification (building on cases described in Higginbotham [2010]), but I’ll grant that there are at least some interesting cases of IEM judgments here.

² While representative of reasons for interest in IEM, this list is by no means exhaustive.
ourselves does not only include the fact that we are thinking things, or bearers of mental properties, as Descartes may have suggested (...) And our first-person conception of ourselves does not only include the fact that we are bearers of physical properties such as being extended in space, as Evans suggests (...) Our first person conception of ourselves also includes the fact that we are the bearers of temporal properties. Our First-person conception of ourselves in other words, is the conception of an object which has a history (Fernández, 2019, pp. 169-70).

If Fernández is right, determining whether memory judgments are immune to error through misidentification should have profound implications, and it is quite understandable why the thesis has enjoyed considerable popularity. Nevertheless, the thought that episodic memory might have some special epistemic status, that misidentifications in it might be impossible, is at least somewhat surprising for the reasons I turn to next.

2. Why episodic memory doesn’t seem to be IEM

Let us adopt Fernández’s *prima facie* plausible functionalist account of episodic memory according to which (episodically) remembering a fact consists in having a mental image that plays the mnemonic role for that fact in the subject.

\[F\]or any subject \(S\) and proposition \(p\), \(S\) remembers that \(p\) just in case \(S\) has some mental image \(i\) such that \(i\) tends to cause in \(S\) a disposition to believe both that \(p\) and that \(S\) experienced that \(p\), and \(i\) tends to be caused in \(S\) by having experienced that \(p\) (Fernández, 2019, p. 49).

Now consider the following fictional case adapted from the cartoon *Bob’s Burgers*.

Aunt Gayle: It was my second sophomore year in junior college, and I took a trip to New York City with...my sister Linda...We were in the denim district...when an elegant woman walked by. I looked up, she looked at me, and she gave me a wink and the finger guns... And that woman was Delta Burke...It was the briefest of gestures, but it meant so much to me. A Designing Woman. That moment helped design this woman.

---

3 It is endorsed, in one form or another, by Bermúdez (2012; 2013), Evans (1982), Hamilton (2007), McCarroll (2018), Recanati (2007), and Shoemaker (1970) among others, though naturally, not necessarily for the reasons Fernández provides.

4 Adapted from *Bob’s Burgers*, Season 7, Episode 22: ‘Into the Mild’ (2017). While the case is fictional, it depicts a familiar phenomenon. See discussions of it and related phenomena in the empirical literature in, e.g. Brown, Croft Caderao, Fields & Marsh (2015); Pasupathi & Wamnryb (2018); Pillemer, Steiner, Kuwabara, Thomsen & Svob, (2015); Reese & Brown (2000); Ross & Ward, (1996) and Sheen, Kemp & Rubin, (2001). Michaelian (2020) briefly discusses such cases but does not pursue them as part of his argument against the claim that episodic memory is IEM on the basis of the existence of observer memory.
Linda: Unbelievable.

Gene: Seriously, it’s hard to believe Aunt Gayle was close enough to have “smelt-a” the Delta.

Linda: Yeah, hard to believe because it didn’t happen to her. It happened to me.

Louise: Wait, Mom, are you saying that Aunt Gayle stole your Delta Burke wink and finger gun story?

Linda: Yeah, it was me who got winked and gunned. We had a connection. I’ll never forget her eyes.

Tina: If it bothers you, you should say something, Mom.

Linda: No, no, it’s fine. It means so much to her, let her have it. I’ve got a good life.

Suppose that Aunt Gayle is sincere in her recounting of the story. She is telling it on the basis of her memory of the trip that she and her sister took, and her belief that Ms. Burke gave her a ‘wink and the finger guns’ is based on a mental image as of it happening just as she described. Such a case satisfies FTM. The subject has a kind of mental image that tends to cause in her a disposition to believe both that \( p \) and that she experienced that \( p \). And such an image is the kind of thing that tends to be caused by such an event happening.\(^5\) In short, Gayle’s mental image plays the mnemonic role for her, and so is an episodic memory that \( p \).

Now, suppose also that Linda is right. While Gayle was present and saw the ‘wink and the finger guns’ happen just as retold, she was wrong about to whom it happened; it happened to Linda. In such a case, Gayle’s belief appears to be an exemplar of a judgment that is subject to error through misidentification. It gets everything right, and is based on a memory that gets everything right, save the identification of who the event happened to (i.e. who exemplified the property in question).\(^6\)

Such errors are familiar and perhaps happen more frequently than we recognize. On reflection, this should not be surprising, given what we know about the mechanisms by which memories are formed, modified over time, and reconstructed (often repeatedly).\(^7\) Memory processes simply aren’t suited to provide immunity to errors in general; identification of subjects in memory would be quite an outlier if it were immune in the way characteristic of IEM.\(^8\)

---

5 This assumes that Gayle is not somehow systematically disposed to delusions. Compare Fernández’s discussion of Korsakoff’s patients (2019, chapter 2).

6 Compare, for example, Wittgenstein’s (1958, pp. 66-7) example of mistaking a neighbor’s broken arm for one’s own on the basis of accurately seeing said arm after a car accident. In that case, the subject would be right in believing that someone’s arm exemplified the property of being broken, but wrong in believing that they were that someone.

7 See, e.g. Schacter & Addis (2007) for a representative model of memory processing.

8 Michaelian (2020) makes effectively the same point (section 1)
Moreover, when you compare beliefs grounded in episodic memory to other judgments that are plausibly IEM, something is crucially different. Here is, for example, Jim Pryor on pain experiences:

> Beliefs of the form *I have a pain* are surely [IEM] when they are believed on the basis of introspective grounds. How could I possibly know on those kinds of grounds that *someone* has a pain, but be wrong in believing that it’s me who feels a pain? *One can’t be introspectively aware of a pain without thereby feeling that pain, oneself* (Pryor, 1999, p. 283; emphasis added).

And here is Giovanni Merlo on the feeling of thirst:

> [I]n judging *that I’m thirsty*, I am not completely sure to be right, but I can ‘rest assured’ that I am not committing an error through misidentification (it would hardly make sense for me to wonder whether I’m mistaking someone else’s thirst for mine) (Merlo, 2017, pp. 613-4), drawing on Wittgenstein (1958, pp. 66-7).

In such cases, and in contrast to ‘ordinary judgments’, there is a sense in which the author of the judgment *can rule out* the possibility that their judgment involves a case of misidentification. Jim can rule out the error because introspectively being aware of a pain involves feeling that pain oneself. Giovanni can rule out the possibility of error because, in making the judgment about the thirst, there is a sense in which it is *nonsensical* for him to wonder whether he’s mistaking whose sensation it is.

Judgments on the basis of episodic memory are not like this. Gayle’s awareness of the wink as happening to her doesn’t make it happen to her and, far from being able to rule out an error of misidentification, wondering whether one’s memory involves such an error is eminently sensible. Nevertheless, Fernández argues that despite such appearances, episodic memory really is IEM, and I turn to this next.

### 3. IEM and mnemonic content

Fernández’s argument hinges on two choices. First, he follows Shoemaker in characterizing IEM in the context of episodic memory as follows:

9 Shoemaker (1968, pp. 563-4) goes so far as to claim that ‘in being aware that one feels pain one is, tautologically, aware not simply that the attribute feels pain is instantiated, but that it is instantiated in oneself.’

10 Merlo (2017, pp. 613-4). Cappelen & Dever suggest that such ruling out could be done *a priori* (2013, pp. 131).

11 Here again, I’m setting aside the argument in Cappelen & Dever (2013, pp. 130-133).

12 Such wondering occupies much of the remainder of the plot of the episode of the cartoon, and is integral to much of the research cited in note 4.
My memory report could of course be mistaken, for one can misremember such incidents, but it could not be the case that I have a full and accurate memory of the past incident but am mistaken in thinking that the person I remember [exemplifying the property] was myself (Shoemaker, 1970, pp. 269-70, emphasis added).

The ‘full and accurate memory’ specification is crucial to Fernández’s understanding of IEM and what it takes for a judgment to fail to be IEM. Here is his definition of IEM (2019, p. 142):

For any property P and grounds G:

If S judges that S has P on the basis of G, then that judgment is IEM relative to G iff it is impossible that there is a subject S* such that:

G represents S* as having P.
G is fully accurate.
S mistakenly thinks that S is identical with S*.
S’s judgment that they have P is false because of (iii).\(^{13}\)

Defined in this way, whether a judgment is IEM or not, is not merely a matter of whether or not it is impossible to form a false belief because of a misidentification of the person or object about whom one made the judgment;\(^{14}\) and a failure of IEM does not necessarily happen when, in forming a judgment that is otherwise correct, someone mistakes something for something else.\(^{15}\) For Fernández, a failure of IEM requires a misidentification on the basis of grounds that are accurate full stop. The result is that whether episodic memory judgments are IEM crucially depends on the nature of the relevant grounds, i.e. on the nature of the content of episodic memory.

The second key choice is Fernández’s decision to adopt a reflexive account of episodic memory content.\(^{16}\) It is as follows:

Reflexive View

For any subject S, memory M and proposition q:

If S has M and S would express M by saying that they remember that q, then there is a perceptual experience P that S would express by saying that they perceive q, such that the content of M is the proposition \{W: In W, M is caused by S having perceived that q through P\}.\(^{17}\)

\(^{13}\) For presentational purposes, I have modified Fernández’s formulation in ways that do not change the account.

\(^{14}\) Pace the earlier mentioned Prosser and Recanati (2012, p. x)

\(^{15}\) Pace Merlo (2017, p. 605)

\(^{16}\) This choice is, of course, eminently understandable, given that he spent chapter 3 motivating said account.

\(^{17}\) Fernández (2019, p. 79)
Defining mnemonic content in this way entails that one’s memory (i.e. ‘G’ in the above formulation of IEM) is only fully accurate if the remembering subject is identical to the person or object about whom the memory judgment is made. Prosser’s formula is thus satisfied as it is impossible for the causal source that leads to the application of the predicate and the object to which the predicate gets applied to fail to be identical. In other words, there is no room for misidentification. So, while subjects may have memory-related judgments that involve misidentification as described in section 2, such judgments do not, technically, fail to be IEM because they are not formed on the basis of grounds that are accurate full stop.

4. Re-examining the significance of IEM

Once one grants Fernández’s definitions of IEM and mnemonic content, it follows, effectively by definition, that episodic memory judgments are immune to error through misidentification.\textsuperscript{18} In this section, I conclude by briefly arguing that this result is not as interesting as Fernández suggests.

First, while it may be that memory judgments are technically IEM, they are still importantly different from other paradigmatic cases of IEM judgments. For example, as we saw above, in judging that one is in pain on the basis of introspection, one’s awareness (and corresponding judgment) of the sensation at that moment involves oneself having it at that moment. It is this intimate connection that purportedly makes it ‘nonsensical’ to wonder whether it is oneself undergoing the experience, and that grants the author of the judgment some kind of epistemic assurance, i.e. that explains its being IEM.\textsuperscript{19} In contrast, on the current proposal, it is the fact that any memory-related judgment that fails to be IEM will also fail to be based on ‘fully accurate’ grounds that explains its being IEM.\textsuperscript{20} Fernández rightly, I think, argues that assessing the IEM status of episodic memory by examining quasi-memories would be unproductive because quasi-memories are defined in such a way that errors through misidentification are possible (2019, Chapter 6, section 3). However, whereas quasi-memories are defined such that errors through misidentification must be possible,
Fernández endorses definitions of mnemonic content and IEM that make (the relevant) misidentifications impossible. While this enables him to draw the desired conclusion, it also rings hollow because it secures his desired conclusion by fiat and thereby trivializes the thesis.21

As a result, and more importantly, it is not clear that it is the fact that episodic memory judgments are IEM that reveals anything about our awareness of ourselves in memory, or our self-conception. On the present proposal, memory judgments are only IEM once one adopts a particular account of mnemonic content (RV). Importantly, according to said account, memories represent, among other things, past experiences had by the remembering subject and causal relations between those past experiences and the remembering subject’s current representations. Thus, it is the account of memory content that delivers insight into how we are presented to ourselves in memory, and the nature of our self-conception as of being the bearers of temporal properties—objects with histories. It is not the fact that memory judgments are IEM. In sum, the means by which Fernández’s proposal secures the claim that episodic memory judgments are IEM strips that quite surprising result, and potentially interesting claim, of its theoretical import.

References


21 A referee helpfully noted that this parallels the circularity of a so-called Lockean criterion for personal identity/persistence (whether Locke endorsed this view is questionable [see Behan [1979] cited in Olson [2019]]. In short, one can secure the result that person A at t2, is identical with person B at t1 only if A at t2, can remember an experience had by B at t1, by assuming that one can only remember one’s own experiences. However, doing so makes the result is trivial. See, e.g. the discussions in Grice (1941, p. 344), Shoemaker (1959, p. 870); (1970, p. 270), and Parfit (1984, pp. 219-23). While Fernández does much to argue for his self-reflexive account of memory content, in the context of a distinct argument for the claim that episodic memory judgments are IEM, adopting the self-reflexive account amounts to begging the question in much the same way that assuming one can only remember one’s own experiences begs the question of personal persistence.


