**Memory without identity**

Abstract: I defend the view that episodic memory judgements do not depend on any kind of identification of oneself as the person whose past is being remembered, and are therefore logically (rather than merely *de facto*) immune from error through misidentification relative to ‘I’. There are two challenges to this view that have been pressed in the literature. One appeals to the idea of background presuppositions of identity and says that ‘I am the person from whom my memory impression derives’ is a background presupposition of any memory judgment. The other appeals to wh-misidentification and says there are possible cases in which memory goes astray that should be counted as cases of error through wh-misidentification. Although the details are different, the core thought behind both challenges is the same: Shoemaker’s thought that there could be a memory-like relation that one stands in to another’s past. I think this thought is a mistake.

*Key-words: IEM, memory, background presupposition, epistemic dependence, semantic presupposition,*

Author details:

Daniel Morgan

University of York

Sally Baldwin Buildings Block A, Wentworth Way, Heslington, York YO10 5DD

Danielmorgan1917@gmail.com

Is episodic memory tied to the past of the subject doing the remembering? I can of course remember other subjects’ pasts. I can remember meeting my brother’s dog for the first time, another subject. But in that case what I remember ‘from the inside’ is not *having met a human*, but *having met a dog.* It’s natural to think that when I gain or retain knowledge about someone’s past by having a memory ‘from the inside’ – as I do in this case just in relation to *meeting the dog* – the subject whose past I am thereby in touch with has to be me.

However, Shoemaker noted the in principle, albeit empirically remote possibility, of the following happening: a subject in the past really had a certain property, that subject wasn’t me, their having had that property causes me now to have the apparent memory ‘from the inside’ of having had that property.[[1]](#footnote-1) The exact mechanisms of this could be filled in different ways but one would be to suppose that someone else met the dog, they retain a vivid memory impression of having done so (no doubt with much more determinate content than what I have reported: the impression could be committal on details of the shape, and size, and smell of the dog and of the environment in which the dog was situated); that memory impression is realized in a certain isolable and extractable part of their brain, and then through some kind of surgical intervention the relevant part of their brain is grafted into mine so that I now have the a subjectively similar memory impression to the one they formed. I’ll assume that this is at least a logical possibility. A take Shoemaker had, and others have also had, is that this shows that memory judgments *are* vulnerable to error through misidentification relative to ‘I’, even if *de facto* – as a matter of contingent empirical fact– such errors never occur. Shoemaker’s idea is that were this kind of error to occur I would make a mistake in *identifying* the one who met the dog as myself. So, memory judgments are not ‘logically’ or ‘*de jure’* IEM.[[2]](#footnote-2)

Two ways of vindicating this take have been pressed in the literature. The first appeals to the idea of identity propositions as *background presuppositions*. Background presuppositions are propositions that a claim epistemically depends on, even though they are not included in the subject’s grounds for the claim. The idea is that Shoemaker’s thought-experiment shows that memory judgments like ‘I met a dog’, have the identity ‘I am the subject from whose past my memory impression derives’, as a background presupposition, and therefore epistemically depend on that identity.[[3]](#footnote-3) The second approach is to appeal to the notion of wh-misidentification. The idea is that there’s a kind of support that memory provides for ‘Someone met a dog’ that is independent of the support it provides for ‘I met a dog’ and in a case like the one Shoemaker imagines I might achieve knowledge of the existentially proposition about *the past* even thought I don’t achieve knowledge of the singular judgment about *my past.*[[4]](#footnote-4)

There’s been a certain amount of dispute between theorists who use the first of these routes to articulate the significance of Shoemaker’s kind of case, vs those who use the second. For example, Coliva, who champions the first of them, is skeptical about the second (see 2024:). But the two routes do both have a great deal in common – they are both trying to articulate the significance of Shoemaker’s case. So we would expect that that there would be something in common with the diagnosis of where the two attempts go wrong, if they go wrong.

The same case figures in my diagnosis of both attempts. It’s a case where someone judges, on the basis of introspection, that they are in pain. In fact, they merely have an itch, not a pain, but their having this itch is caused, by someone else being in pain.[[5]](#footnote-5) My idea is that thinking about this thought experiment side by side with the Shoemaker thought experiment helps illuminate the latter.

The pain-itch case is apt to seem rather trivial, showing nothing interesting about the epistemic architecture of introspective judgments. I totally agree. In fact I take as a premise that the pain-itch case should not be taken to show that introspective judgments are vulnerable to error through misidentification relative to ‘I’. The idea is to use the side by side comparison to show that the same is true of Shoemaker’s case. Regarding the two types of mistake, it’s a levelling-down exercise rather than a levelling up exercise: neither case turns out to show anything significant. But regarding the two faculties involved in the two cases – memory, introspection – it’s a levelling up exercise. I think that introspective judgments like ‘I am in pain’, and memory judgments like ‘I was on a ship’ are both immune to error through misidentification, logically and not just *de facto*.

Of course, my argumentative focus is primarily negative. Memory judgements seem to be IEM. In this paper, I rebut the two main ways in the literature of arguing that in fact they aren’t IEM. But, of course, there could in principle be some third, entirely different way of arguing that they aren’t that has not occurred to anyone yet. If you’re convinced that memory judgments can’t be IEM, then the upshot of this paper is just that you have work to do in articulating the notion of IEM that they don’t satisfy, since neither of the existing two strategies for articulating it work. But, I submit, it would also be reasonable to revisit the question of whether there is any good reason to think that memory judgments can’t be just what they seem to be: IEM.

Part 1 says why I think the background presupposition route to crafting the required notion of IEM does not work. This draws on a style of argument that Morgan 2019 utilises but that has not been convincingly replied to. I suggest that the original argument can be strengthened by thinking about semantic (not epistemic) presupposition. Part 2 focuses on the challenge from wh-EM, where I think what I previously argued for is not fully convincing, for reasons McGlynn 2021 brings out. I try to give a better argument for the claim that wh-IEM does not vindicate Shoemaker’s position. In concluding, I touch on the question of whether Evans, whose skepticism about the significance of Shoemaker’s cases I am attempting to vindicate, was also right to think that the debate about which judgments are IEM relative to ‘I’ was relevant to the correct metaphysics of the self.

This is going to be quite a literature-heavy discussion. In particular, I’m going to refer quite a bit to the incredibly helpful introduction to this volume provided by Coliva and Palmira (2024). In approaching the topic in this way, I’m moved by a worry that John Schwenkler expressed in a review of a previous collection of essays on immunity to error through misidentification:

the contributors generally proceed in isolation from one another, even when their positions conflict. Especially given the frequent disagreements and the difficulty of seeing any way to resolve them, one general worry that emerges from these diverse treatments is that there may be no unitary phenomenon of IEM for philosophers to understand, as the concept of IEM encompasses a group of phenomena that have some features in common but are insufficiently homogeneous to work cooperatively in any general account of self-consciousness or singular thought. (2013: 180)

Coliva and Palmira’s discussion has helped me get on top of what the main challenges to my view about memory is. It’s also helped me understand that that view and another thesis I’ve also defended but won’t discuss here, the *Simple Explanation*, are mutually reinforcing, in ways I hadn’t appreciated. [[6]](#footnote-6) So in the spirit of cooperation I try to use the introduction as a helpful herding device for my own thoughts, and hopefully that also does something to address Schwenkler’s very reasonable worry about contributions to the IEM being too disparate.

## Part 1: Identities as background presuppositions

What do cases like Shoemaker’s show about memory? One take, which Shoemaker endorsed and Evans denied, is that they show that such judgments are always *based on* or *grounded in*, the following kind of identity:

I = the person whose past is responsible for the memory impressions I am having.

A popular thought is that Shoemaker and Evans were both onto something. I quote Coliva and Palmira’s description of Coliva’s (and García-Carpintero, Wright’s and Peacocke’s) way of trying to capture what both got right:

Coliva has then proposed a different resolution of the Evans-Shoemaker dispute, based on the distinction between a subject’s available grounds for their judgement and a judgement’s background presuppositions. Let us take memory-based self-ascription ‘I was in Scotland five years ago’. As persuasively argued by Evans (1982), one (1) neither arrives at such a self-ascription *by going through a piece of conscious reasoning* that takes identity belief “I am the person from whose past this memory impression derives” as a premise, (2) *nor would one offer such a belief as an explanation* of why one is making that self-ascription. For, if one were asked how one arrived at self-ascription, one would most likely only cite one’s memory experiences. This has led Coliva and many others (see e.g. García-Carpintero 2018, Wright 2012) to say that the identity belief “I am the person from whose past this memory impression derives” is not foregrounded in one’s own psychology and, for this reason, is not part of the grounds of one’s self-ascription. However, that identify belief is a background presupposition of one’s self-ascription. (3) *If (one knew that) the identity belief were not in place, then one couldn’t rationally hold “I was in Scotland five years ago”.* (2024: my italics, and numbering)

When one judges, one’s judgment may epistemically depend on other claims. The idea is that epistemic dependence comes in two species. One species of epistemic dependence is the kind of epistemic dependence a judgment has on its *grounds*. The other species is the kind of epistemic dependence a judgment has on its *background presuppositions*. Evans was right that memory judgments are not *grounded* in identities. Shoemaker, the claim is, would have been right if he’d claimed that they nevertheless have identities as *background presuppositions*.

It will be important how we test for these different kinds of epistemic dependence. I begin with grounds, and how we test for them. The rough idea is that *x* is part of your grounds for a judgment if and only if: either *x* figured in the conscious reasoning that led you to *j or* it didn’t but, if *j* were challenged, you would appeal to *x*. These disjuncts are adverted to by (1) and (2) respectively. (1) and (2) together imply that memory judgments like ‘I was in Scotland’/‘I was on a ship’ do not have ‘I am the person from whose past this memory impression derives’ as a ground.

This test for grounds is arguably only rough, because of a point that Wright makes:

You ask me, “Why do you judge that your Aunt Lilian is wearing an extraordinary hat today?” and I might reply: “Well, this is she—this woman is my Aunt Lilian—and, as you can see, she is indeed wearing the most extraordinary hat”. Of course there is some murkiness about this. The notion we want is that of the grounds on which a thinker *actually bases* a given judgment, even if she doesn’t think them through in a fully explicit fashion, rather than of things she might, as it were extemporaneously, say in support of it if pressed—which may of course encompass a much wider class of considerations, once she has the opportunity to re-marshal her thoughts and her information (2012: 268).

This roughness doesn’t do anything to block the move from (1) and (2) to memory judgments not having an identity in their grounds. Wright’s point is that the test could count *too much* (extra things one might come up with if one has time to alight on new considerations), not *too little*, as among the grounds of a judgment.

But Wright’s test is a reminder that it can be quite tricky to test for epistemic dependence relations. And that is relevant for my argument that memory judgments do not have an identity in their background presuppositions. I agree that they pass the counterfactual test Coliva and others have articulated for doing so, adverted to in (3). I agree that there are cases where that test works as a test for background presuppositions. But I claim that there’s something special going on in the case we’re interested in that means the test does not work in that case.

For a case where the counterfactual test for background presupposition works, I continue with Wright:

Suppose I look out of the window and, so it seems to me, see that jay once again in the crab apple tree. And consider any view in the broad spectrum of possible views that agree in allowing that my judgement, [*There is a jay in the crab apple tree*], is one for which my experience provides justification. On any such view, the question arises: what if any part in the justificational architecture of my belief is played by certain collateral but obviously relevant-seeming propositions like that my visual system is currently working effectively, or indeed that the orchard is currently visible through the window — that I am not looking at a clever *trompe l’oeil* painted upon the glass? These are propositions that I do, no doubt, believe; and if I doubted them, I would doubt that my experience justified my belief about the jay (2012: 268).

Wright’s claims, plausibly, that the judgment *there is a jay in the crab apple tree* has as a background presupposition *my visual system is currently working effectively*. The test he appeals to in defending that verdict is a counterfactual test that *reverses* the counterfactual test *for grounds*. The test for whether x is part of the grounds of an ultimate conclusion *j* is whether, if *j* were challenged, *x* is something the subject would think to appeal to. The test for whether *j* has *x* as a background presupposition is whether if *x* were challenged, the subject would still be justified in holding onto *j*. It’s a little bit like the distinction between recall and recognition memory. The more demanding counterfactual is checking whether the relevance of *x* to j is something I’m on top of enough to bring *x* up myself if *j* is challenged (analogy: can I recall what the capital of Tuvalu is). The less demanding one is checking if whether if someone specifically brings up the issue of whether x is relevant to j, I can at least recognize that it is relevant (analogy: can I recognize as correct the claim that Funafuti is the capital of Tuvalu). Our recall memory far outstrips our recollection memory, and arguably the claims our judgments epistemically depend on far outstrips the propositions they are grounded in.

Assuming that the counterfactual test for background presupposition works, the key point is that:

*If ‘I am the person from whom my memory impression derives’ were not in place, then ‘I was on a ship’ would not be justified.*(3)

In 2012, Coliva talks about claims ‘falling into question’ rather than not being ‘in place’, and what a rational and fully conceptually equipped subject would do in that case. I’ll base my discussion around that phrasing.

Suppose I judge, on the basis of a memory impression, ‘I was on a ship’. Someone comes along and queries whether, as they would put it, ‘You are the person from whose past your memory impression derives’. Equipped as I am with the concepts they are deploying, I understand them. Rational as I am, I will see the relevance of what they say to my ‘I was on a ship’ judgment, based on memory. So if their query leads the identity to fall into question for me, I surely will retreat from ‘I was on a ship’.

I deny that if follows that the memory judgment epistemically depends on an identity by having an identity as a background presupposition. First, I show *that* the counterfactual diagnostic *over generates* epistemic dependence. Second, and more tentatively, I offer a proposal about *why and when* it over generates epistemic dependence.

For ‘*that*’, take an arbitrary singular judgment, ‘*a* is F’. Now consider a scenario in which it is called into question whether *a* is the entity whose being F is responsible for my judging that *a* is F. In that scenario, I would of course come to doubt whether I am justified in judging ‘*a* is F’. It follows that ‘a is the entity whose being F is responsible for my judging that *a* is F’, an identity, is a background presupposition of my judgment. Since the singular judgment ‘a is F’ was arbitrarily chosen it follows that, on this definition, every singular judgment depends on such an identity. For example, ‘I am in pain’ based on introspection, would. But ‘a is the entity whose being F is responsible for me judging that a is F’ does not seem to have a similar status to ‘My sense organs are working effectively’. It doesn’t seem like something every ‘a is F’ judgment epistemically depends on.

I turn to ‘*why and when’*. What’s the difference between ‘My visual system is working effectively’ or ‘my memory is and has been working effectively’, on the one hand, and ‘I am the person from whom my memory impression derives’, on the other? I bring this out by pursuing an analogy.

The Warren report, the official report on JFK’s assassination, makes a number of different claims about the event. It characterizes the person it says was the shooter (Oswald) as an unconnected nobody, not an agent of any intelligence service or crime syndicate. It also claims that there was exactly one shooter of JFK. Suppose I start out accepting the Warren Commission’s line on everything, and that it’s reasonable for me to do this. So, in particular, I start out with overall grounds for believing:

 The shooter of JFK was an unconnected nobody.

Someone, *Critic*, who it’s highly reasonable for me to trust, including when what they say diverges from the Warren Commission, asserts:

 The person who shot JFK was really an agent of the CIA.

As a result, my grounds no longer provide overall support to the shooter having been an unconnected nobody, and I abandon that characterisation in favour of *Critic’s* rivalcharacterisation.

Notice that I still have some grounds to believe somethings the Warren Commission says. For example, I still have grounds to believe:

 There was exactly one shooter of JFK.

Why do I end up with grounds to believe this?

One explanation is that this is something the Warren Commission says that *Critic* didn’t specifically target, and I retain the grounds Warren Commission provides for any proposition that isn’t specifically targeted by *Critic.*

It’s not obvious though that I do retain the grounds the Warren Commission provides for any proposition that isn’t specifically targeted by *Critic*. Maybe if I can’t trust the Warren Commission on such a basic claim about what sort of person the shooter was, I can’t really trust it on anything. But, whatever about that, appealing to my grounds *deriving from the Warren Commission* to believe that there was exactly one shooter misses something out. What it misses out is that I now have *Critic’s* word for it that there was exactly one shooter. *Critic’s* testimony *gives me new grounds* to believe that. *Critic* didn’t *assert* this proposition. But they *semantically presupposed* it*.*

The source of the semantic presupposition is the definite description ‘the shooter’. Claims of the form ‘The F was G’ entail that there was exactly one F. Because this entailment is a semantic presupposition, *negating* the sentence (‘It is not the case that the shooter of JFK was connected to the CIA’) does not make a difference to the entailment. Similarly, the entailment projects into contexts involving attitude verbs or speech act verbs. Consider:

*Critic* called into question whether the shooter of JFK was an unconnected nobody.

This longer sentence still semantically entails that there was exactly one shooter of JFK.

I now try to say what the analogy is between *Critic’s* intervention and the impact on my evidential situation of it’s being called into question whether I am the person from whose past my memory impression derives.

I start out with a memory impression as of having been on a ship and so with memory grounds for believing:

 I was on a ship.

Now along comes along someone I should trust who calls into question whether *I am the person from whom my memory impression derives*. Suppose they utter the sentence ‘You are not the person from whose past your memory impression derives’. This, let us suppose, means that I end up without overall grounds to believe ‘I was on a ship’. But I do end up with grounds – new grounds, the grounds provided by the person who has queried the accuracy of my memory – to believe:

 There is exactly one person from whose past my memory impression derives.

The impression of someone having been on a ship that derives from this other person’s past could be inaccurate through ordinary non-identity involving mistakes (e.g. maybe it was really just a boat that they were on, not on a ship), just as a memory impression can be inaccurate through ordinary non-identity involving mistakes. But we normally think you can know, on the basis of memory, despite the possibility of this kind of ordinary mistake. So we can set this kind of mistake aside. Having done so, there is at least some pull towards the idea that you end up in a position to know:

 The person from whom my memory impression derives was on a ship.

But you aren’t in a position to know this *just on the basis of memory*. You have mixed *memory-and-testimonial* grounds for it, where the testimony is coming from the person who calls into question whether you are the person from whose past my memory impression derives. And, I say, the fact that the grounds have changed mean that the possibility of ending up in this position doesn’t tell us anything about the epistemic architecture of the original judgment, for which the only grounds were *memory* grounds.

This original grounds/grounds smuggled into a defeater distinction arises recurrently in the literature of IEM. Without it, we risk trivializing the notion of IEM. Aidan McGlynn puts the point very nicely, writing about Joel Smith’s discussion of a different notion of wh-IEM:

If our characterisation… permits the defeating evidence to provide one’s justification for the fallback existential, then whether one’s original evidence in a given case would be capable of supporting that existential claim becomes irrelevant to whether the characterisation is met. So the proposal Smith floats here looks well motivated; we should require that one’s *original evidence*, and not merely one’s defeating evidence, can justify the fallback existential in the face of evidence that defeats one’s de re judgment. (2021: 2297)

I am saying that passing the counterfactual test Coliva and Wright endorse is not enough to show that a judgment epistemically depends on an identity. And that’s because the relevant counterfactual does not respect the ‘original grounds’ condition Smith and McGlynn both rightly insist on, on pain of trivializing the notion. What makes the smuggling in of the new information something that can easily escape notice is that things semantically presupposed register less obviously than things outright asserted.

My take is the counterfactual test for background presupposition is only effective in cases whether the antecedent of the counterfactual doesn’t involve the subject getting *new* grounds for a relevant proposition.

(A)If it were called into question whether *my* *memory was functioning effectively*, a rational subject would retreat from ‘I was on a ship’.

(B)It if were called into question whether *I am the person from whose past my memory impression derives*, a rational subject would retreat from I was on a ship.

I say the truth of (A) is relevant to what propositions memory judgments epistemically depend on, but the truth of (B) is not.

I want to consider two replies to my argument regarding why the counterfactual test does not work in this case.

One is to stipulate that ‘calling into question “You are the person from your memory impression derives”’ shouldn’t be understood the way I have understood it, to say that it’s instead meant to be understood as calling into question whether the impression derives from you, but leaving it open that it may not derive from *anyone at all*.

My reply is that isn’t the kind of thing that can be stipulated, since it clashes with the fact that ‘the’ semantically presupposes uniqueness.

The other reply I want to consider is that ‘The person from whose past my memory impression derives was on a ship’ is not a relevant proposition. That is, all that’s being claimed is that if ‘The person from who your memory impression derives was not on a ship’ were called into question, you would have to retreat from ‘I was on a ship’. It’s not being claimed that you would in that case be in a position to know ‘The person from whose past this memory impression derives was on a ship’. And so it’s not relevant that, if you did end up in a position to know this, the grounds on which you knew it would be *new* grounds, compared to the ground you had before the intervention.

If that is the reply, I fall back on the original overgeneration point involving the pain-itch case indicating that the counterfactual test over generates epistemic dependence. But I submit that without its being claimed that, with the identity called into question, you still end up in a position to know ‘The person from whose past my memory impression derives was on a ship’, we lose the idea that we have an error *through misidentification* here. An error through misidentification sounds like a case whether, at least as regards the predicative part of the judgment, things are epistemically OK. But if someone pursuing the background presupposition based strategy for finding error through identification is unwilling to claim that the subject ends up in a position to know anything involving the relevant predication (‘was on a ship’) we’ve lost the idea that this is a kind of error through misidentification. On the other hand, if their claim is that you do end up in a position to know ‘The person from whose past my memory impression derives was on a ship’, the point that that is only because of the new grounds that are smuggled into the defeater kicks in.

### Part 2: Wh-misidentification and Shoemaker’s challenge

‘IEM relative to background presupposition’ is one kind precisification of IEM that has been taken to show that memory judgments are vulnerable to misidentification. Pryor introduced the other kind, the concept of wh-IEM. He argued that memory judgments were not wh-IEM, because of Shoemaker’s type of case.

Wh-EM does not look as immediate a threat to memory judgments being IEM as ‘EM relative to background presupposition’. With background presupposition-EM, the impetus for recognizing the relevant notion of EM and for thinking that it’s a notion of EM that memory judgments are vulnerable to it are one and the same: Shoemaker’s thought-experiment. With wh-IEM, Pryor introduced the notion by reference to a case that wasn’t a first-person judgment at all, his skunk case. He argued that memory judgments like ‘I was on a ship’ failed to be wh-IEM. But it wasn’t just obvious, looking at how the notion was introduced, that it would turn out either that this was true, or that there was some problem with the notion of wh-IEM.

I have previously argued that, pace Pryor, memory judgments are wh-IEM. Here, I note a challenge to my argument from McGlynn, concede I don’t have a reply to it, and suggest a modified way of defending the claim that memory judgments are IEM from the wh-EM challenge. The key idea is that there are multiple different notions of wh-EM but none that enable us to distinguish memory judgments from introspective judgments.

How should we define wh-EM? One approach would be to say that the notion is Pryor’s to stipulate the contours of. But Pryor himself didn’t encourage this approach, and certainly nobody else took it. Pryor himself offers his definition of wh-EM, which appeals to the idea of an undercutting defeater, as a substantive story which could turn out to be wrong. Different theorists have said it was wrong, and offered their own rival theories.[[7]](#footnote-7)

If we can’t assume that wh-IEM is just exactly what Pryor says it is, what can we assume? I think there are two minimal assumptions.

First, in *the skunk case* the error that occurs is an error through wh-misidentification. In the skunk case, one has olfactory grounds for judging ‘There is a skunk in my garden’. One has visual grounds for judging of an animal one sees that it is a skunk. And, putting the visual and olfactory grounds together one judges ‘That is a skunk in my garden’. If this is wrong, e.g. because the animal isn’t a skunk and the real source of the odour is a skunk hiding behind this animal, then this, stipulatively, is an wh-EM.

Second, vulnerability to wh-misidentification should have something to do with the distinction between judgments of the form *a is F* and *Something is F* and of the possibility of scenarios in which the latter retains some relevant status (e.g. justification, truth, knowledge…) and the former lacks some relevant status (e.g. justification, truth, knowledge…).

 The account of wh-IEMI defended was as follows:

A judgment of the form ‘a is F’ is, relative to certain grounds g, immune to error through wh-misidentification if and only if g justifies ‘a is F’ without offering independent knowledge of ‘Something is F’.

I offered the following test for independent knowledge:

grounds g offer knowledge of q that is independent of the justification they provide for p if and only if one can know that q on the basis of g, even if p is false.

For example, even if it is false that that animal you can see is a skunk one can know ‘There is a skunk in my garden’ just on the basis of how things smell. By contrast, I claimed, if one wasn’t on a ship, because one is in Shoemaker’s case, one can’t know ‘Someone was on a ship’, just on the basis of one’s memory impression. Hence, smell provides offers independent knowledge of ‘There is a skunk in my garden’ but memory does not provide independent knowledge of ‘Someone was on a ship’. Hence, ‘That is a skunk in my garden’ is vulnerable to wh-EM relative to ‘that’, but ‘I was on a ship’ is immune to wh-EM relative to ‘I’.

I said what makes the test appropriate is a ‘fact about knowledge’, that one cannot gain knowledge from false premises. So, if one’s knows ‘There is a skunk in my garden’ in this case one can’t knowing it by inferring it from ‘That is a skunk in my garden’, since that is false. One must therefore have some independent way of knowing ‘There is a skunk in my garden’. And one obviously does. One smells a skunky odour and that justifies the existential generalization all on its own.

But McGlynn noted a problem:

There have been a number of examples that suggest that the ‘fact about knowledge’ Morgan appeals to, that there’s no getting knowledge from a falsehood, is simply false. Take an example from Ted Warfield (2005). I might have miscounted the number of people at a talk I’m giving, reaching a total of 67 when they are actually 66 in the room. If I infer from this that the 100 handouts I’ve printed off will suffice, this seems to be knowledge even though my premise is false. My purpose here is not to join or try to resolve this debate about knowledge (though I do think that knowledge from falsehood is possible: see McGlynn 2014: 7). My point is just that there’s a major epistemological controversy at the heart of Morgan’s account (2021:2310).

One reply to Warfield’s putative counterexample is to say that this is knowledge *despite* falsehood rather than knowledge *through* falsehood. I don’t know ‘There are 67 people in the room’, since that is false. But, on the basis of counting, I know something like ‘There are *roughly* 67 people in the room’ and it’s from this true premise that I infer, and thereby come to know, that there are fewer than a hundred people in the room.

A further point we can make is that, as this case is naturally imagined, the agent also has a *perceptual* justification for believing ‘There are fewer than 100 people in the room’ that is independent of counting and inference. One can get a rough sense of numbers just by looking out at a room. One has this independent perceptual route to knowledge of rough number even in cases where one has, additionally, done a count. We’re being instructed by Warfield to ignore this perceptual route to knowledge (since we’re being told to focus on inferential knowledge). But just as one might wonder how safe it is to stipulate that the inference one performs is from ‘There is 67 people in the room’ (as opposed to ‘There are *roughly* 67 people in the room’) one might wonder how safe it is to stipulate that the judgment we’re interested in is based on counting and inference (as opposed to based on *perception*). In both cases, one might think that our intuitive verdict about the case – that you end up knowing roughly how many people there are – could be based on imagining the case in a way that clashes with the stipulation. That is, we imagine that your knowledge comes via inference from the true premise about the approximate number. Or we imagine that your knowledge comes via perception rather than counting and inference. If so, it doesn’t show that you can gain knowledge by inferring from a false premise.

I don’t expect these points to persuade everybody. So I won’t rely on them. I assume in what follows that one can have a counting-and-inference based knowledge of ‘There are fewer than 100 people in the room’, even if the judgment from which one infers one’s conclusion, ‘There is 67 people in the room’, is false.

McGlynn brings up the Warfield case just as a counterexample to a general claim about knowledge that I rely on in arguing for my definition of wh-EM. He doesn’t say whether he thinks Warfield’s case is itself involves vulnerability to wh-EM. But, in fact, it’s interesting to note that question does arise. Although it is a bit forced, we can make the counting case more easily comparable with the skunk case by framing the two quantity judgments as follows:

*67* is the number lower than 100 that is a number of people in the room

*There is* a number of people lower than 100 that is a number of people in the room.

Should we say that the claim that 67 is the number of people less than 100 that is the number of people in the room involves error through wh*-*misidentification relative to ‘67’?

On the one hand, the grounds one has in this case – having done a count– do not decompose in a way that resembles the way that the grounds in the skunk case decompose – seeing, and smelling. If one’s grounds are *having counted* it’s not clear we can make sense of dividing this into parts, one part of which is especially relevant to the existentially general claim. On the other hand, it does seem like one can gain knowledge of the existentially general claim that is independent of the singular claim being knowledge, or even true. If you were told ‘There aren’t in fact 67 people’, it wouldn’t be a shot in the dark to retreat to ‘Well, at any rate, there are fewer than 100’.[[8]](#footnote-8)

I take it that part of what’s controlling our verdict that the existentially general judgment is knowledge in this case is that a capacity to be out by 1 or 2 or 3 on a count is no real evidence of capacity to be out by more than 20, still less that this unfortunate capacity will have been exercised on this occasion. The kind of thing that needs to occur for the singular judgment to be false (the kind of thing that *has* in fact occurred) is very different from the kind of that would need to occur for the existential judgment to be false. So, this rougher judgment semes very reliable. The minor falsity of the premise (‘67 is the number under 100…’) from which the subject derives their conclusion is more than compensated by the weakening that occurs as they move from premise to conclusion (‘There is a number under 100’). A subject who is a highly unreliable detector of the *exact number of the thing being counted* might yet be a highly reliable detector of the *exact number plus or minus 20*. If their approach to detecting the thing they are reliable at is indirect – they first form a judgment about what the exact number is and then they infer the further weaker claim that it is within a range defined by 20 above and below the exact number – that doesn’t make them any less reliable on the issue of what the *exact number plus or minus 20* is. So, we might say that their ‘67’ judgment is vulnerable to wh-misidentification, and, in Warfield’s case, an error through wh-misidentification has occurred. They just make a mistake in identifying which number under 100 was the number of people in the audience.

We can imagine a case involving memory judgments built on these lines. E.g. the subject judges ‘I was on a ship’ based on memory. In fact, this is false, because they are in a Shoemaker type scenario. Moreover, the only kind of mistake they could easily make is a mistake of the sort they do make. There are no nearby worlds in which they think they were on a ship, e.g. simply because they confused a dream about a ship for a memory about a ship. So, they are a reliable detector of *someone having been on a ship* in a way that they just aren’t of whether *they were on a ship.* We might then say that this shows that their ‘I was on a ship’ judgment is vulnerable to wh-misidentification. They just make a mistake in identifying which person was on a ship.

We can equally imagine a case involving introspective pain-judgments build on the same lines. The subject judges ‘I am in pain’, based on introspection. In fact, this is false, because they are in as scenario in which they mistake an itch for a pain. Moreover, the only kind of mistake they could easily make is similar. There are no nearby worlds in which they think they are in pain, e.g. because they mistake an itch for a pain, and the itch is not caused by somebody else’s pain. So, they are a reliable detector of *someone being in pain* in a way that they just aren’t of whether they are in pain. If we were persuaded by the previous paragraph, to say that the ‘I was on a ship’ is vulnerable to wh-misidentification we should say the same about the ‘I am in pain’ judgment. The person just make a mistake in identifying which person was in pain.

We might resist both verdicts about the first-person cases. A relevant difference between the skunk case and the counting case, on the one hand, and the two first-person cases on the other, is that in the first two cases the existentially general judgment isn’t just *reliable*, even across cases in which the singular judgement is false, it’s one that is reasonable *from the subject’s own perspective*, even in such cases. The person in the garden with the skunk understands that vision and smell are separate and that their singular judgments depends on both but the existential judgment only on smell. The person counting understands that, although their grounds don’t decompose into two parts, a minor miscount is much more on the cards than a major miscount. Evans’s talk of someone who is told that their memory judgment is false retreating to the existential generalization being ‘a shot in the dark’ might connote unreliability/flukiness, or unreliability/flukiness *from the subject’s own perspective*. If we think reasonableness of the retreat to the existential generalisation *from the subject’s own perspective* is crucial to ‘That animal is a skunk’ and ‘There is a number lower than 100 that is the number of people in the room’ counting as vulnerable to wh-misidentification, then,we won’t think of the errors involving ‘I was on a ship’ or ‘I am in pain’ as counting as errors through wh-misidentifcation. But, of course, we can change the details of the first-person cases to make them more resemble the skunk case or the counting case. We can stipulate that the person is aware of the curious pattern of vulnerability that their memory judgments have. Equally, we can stipulate that the person is aware of the curious pattern of vulnerability that their introspective judgments have.

I think that if we start out with wh-IEM as (i) stipulatively returning a certain verdict in the skunk case and (ii) stipulatively tied to the possibility of the status of ‘Something is F’ coming apart from the status of ‘a is F’ and (iii) with nothing else stipulated, then there will be different notions that satisfy the stipulations and are to that extent equally ‘correct’ definitions, albeit of different epistemic statuses. These could differ from each other in regard to:

* Whether the grounds *g* for ‘a if F’ need to decompose in the way that the olfactory-visual grounds decompose (and that memory grounds, and counting grounds do not)
* Whether the status of belief in the existential generalization on the basis of *g*, even across cases in which the singular judgment is false, needs to (i) be reliable or (ii) be reliable from the subject’s own perspective or (iii) be both (i) and (ii) or (iv) have some other epistemic status entirely.

I don’t really see how there could be a fact of the matter about which of these is *the* notion of wh-IEM. We can refer to whichever status we’re most interested in. I’ve tried to argue though that none of them will provide a way of arguing that memory judgments are vulnerable to error through wh-misidentification, without tipping into establishing the same thing about the introspective judgments.

## CONCLUSION

I’ve argued that memory judgments are not vulnerable to errors through misidentification. In closing, I note two lacuna in my discussion. One is the question of why this conclusion is significant. Another is why I’ve been assuming that the pain-itch case couldn’t itself be an error through misidentification. I try in the conclusion to say something about both worries at once.

One can remember past bodily properties (e.g. having been on a ship [which is bodily in so far as it presupposes spatial location], having been doubled over) as well as past mental properties (having been feeling seasick). So, it’s not immediately obvious how the putative identity-dependence of memory, i.e. the issue Shoemaker most directly raises, might be relevant to any epistemic asymmetry in access to one’s mental properties vs bodily properties. However, the people (e.g. Coliva, Wright, García-Carpintero) who think that memory judgments are vulnerable to errors through misidentification standardly think that the same is true, and can be shown to be true using similar thought-experiments, of one’s other distinctive ways of gaining knowledge about one’s own bodily properties, e.g. one’s way of gaining knowledge of the fact that one is currently doubled over. And they standardly think that is not true of, for example, one’s special way of gaining knowledge of one’s own mental properties (e.g. that one is in pain). Assuming they are right, the putative identity dependence of memory, can serve a stand in for identity dependence of every way of gaining knowledge of oneself that is, at least in part, way of gaining knowledge of oneself qua bodily subject, qua human animal, and it thereby shows that there is an epistemological asymmetry between our access to our bodily, and mental, properties – only the latter is IEM. One short justification for assuming that that the pain-itch case couldn’t itself be a case of misidentification is that, if it were, showing that memory judgments are vulnerable to errors through misidentification wouldn’t support the epistemological asymmetry claim it is supposed to.

Moving to the second lacuna, why is the epistemological asymmetry claim significant. Coliva’s take on it as follows:

According to Coliva (2012, 2017), the distinction between de facto and logical IEM could give rise to further illusions of transcendence. For it would turn out that only non-inferential psychological self-ascriptions are logically IEM. This, in turn, could give rise to the idea that the self is most basically presented to oneself as the thinker of a given thought and this might be taken to show that it is identical to a Cartesian ego (or to a “stretch” of one). *Yet, this would be a non-sequitur.* For from the fact that one is presented to oneself as a thinker of a thought *it does not follow* that only a mental entity can have introspectively available thoughts. That is, being presented to oneself as the thinker of a given thought is entirely compatible with the fact that one is a physical entity capable of entertaining thoughts and of being aware of oneself while doing so. According to Coliva (2012) while IEM and logical IEM can account for the illusion of transcendence, they are in fact compatible with an animalist conception of the self once these phenomena are correctly characterized. Contrary to Evans and McDowell, then, it is not necessary to deny such a distinction to maintain the view that the self is identical to a living human being with both physical and psychological properties (my italics).

The ‘This would be a non sequitur’ claim, understanding as a claim about deductive validity, seems absolutely right. But it’s not obvious that it’s conclusive. Consistent with it, there seems to be a reasonable abductive argument that starts with the premise that there is an epistemological asymmetry along mental-bodily lines, proposes that metaphysical asymmetry along the same lines would be one adequate explanation for it, and presses the person who thinks that the self is identical to a living human being for what their alternative explanation for it is. An advantage of my view that memory judgments are IEM, if it can be made out, it is that it does not leave us, even if we are animalists, hostage to the task of coming up with an alternative explanation.

**Bibliography**

Coliva, A. (2006) “Error Through Misidentification: Some Varieties”, *Journal of Philosophy* 103(8): 403-425.

Coliva, A. (2012) “Which ‘Key to All Mythologies’ About the Self? A Note on Where the Illusions of Transcendence Come From and How to Resist Them”, in Prosser, S. and Recanati, F. (eds.) *Immunity to Error Through Misidentification. New Essays* (pp. 22-45), Cambridge: Cambridge University Press.

Annalisa Coliva & Michele Palmira (2024), Immunity to Error Through Misidentification: Some Trends

Evans, G. (1982) *The Varieties of Reference*, Oxford: Clarendon.

García-Carpintero, Manuel (2018). ‘De se thoughts and immunity to error through misidentification’. *Synthese.* 195 (8):3311-3333 (2018)

Merlo, Giovanni (2017). Three Questions about immunity to error through misidentification . *Erkenntnis* 82 (3):603-623.

McGlynn, A. (2016) “Immunity to Error Through Misidentification and the Epistemology of *De Se* Thought”, in García-Carpintero, M. and Torre, S. (eds.) *About Oneself* (pp. 25-55) Oxford: Oxford University Press.

McGlynn, A. (2021) “Immunity to Wh-Misidentification”, *Synthese* 199: 2293-2313.

Morgan, D. (2019) “Thinking about the Body as Subject”, *Canadian Journal of Philosophy* 49(4): 435-57.

Pryor, J. (1999) “Immunity to Error Through Misidentification”, *Philosophical Topics*, 26: 271-304.

Prosser, S. and Recanati, F. (eds., 2012) *Immunity to Error Through Misidentification. New Essays*, Cambridge: CUP.

Schwenkler, J. Review of Immunity to error through misidentification (Prosser and Recanati) *Analysi* 73 (1):180-182 (2013)

Shoemaker, S. (1968) “Self-Reference and Self-Awareness”, *The 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 (2):273-83 (2006)

Wright, C. (2004) “On Epistemic Entitlement: Warrant for Nothing (And Foundations for Free?)”, *Aristotelian Society Suppl. Vol.* 78: 167-212.

Wright, C. (2012) “Reflections on François Recanati’s ‘Immunity to Error Through Misidentification: What It Is and Where It Comes From’”, in Prosser, S. and Recanati, F. (eds.), pp. 247-280.

1. See Shoemaker 1968 and 1970. [↑](#footnote-ref-1)
2. It may be helpful to provide a clarificatory note on ‘logically’. One might think that the logically necessary truths are a *subset* of the necessary truths and that ‘Superman is Superman’ is logically necessary, because it’s necessary *in virtue of its logical form* in a way that ‘Superman is Clark’ is not. In the debate about IEM, the point of ‘logically’ is different. It’s not to point at a particular *explanation* of why a judgment is IEM, e.g. to say that it is IEM in virtue of its logical form. It’s essentially contrastive, pointing towards ‘*de facto’*.

How then does ‘de facto’ work? If a judgment is merely *de facto* IEM that means that itis not IEM (it is VEM), but the errors of misidentification in virtue of which it is not IEM only occur in worlds very different from ours. Saying that a judgement is *de facto* IEMis thus more like saying that the Moon is *de facto* immune from the spread of C-19 (which does not imply that the moon is immune from the spread of C-19) than like saying that someone is *de facto* in charge (which does imply that they are in charge).‘*De facto* IEM’ means something like *nearly IEM* or *ignoring far-flung cases IEM*.

 Someone might prefer to say that ‘*de facto* IEM’ really is a kind of IEM, and ‘logical IEM’ is just another more demanding kind of IEM. This wouldn’t affect my argument, which will focus on how we test for the higher status, ‘logical IEM’, whether or not we think that ‘logical IEM’ and ‘IEM’ pick out the same status.

 [↑](#footnote-ref-2)
3. See, for example, Coliva 2006, García-Carpintero 2018, and Wright 2012. [↑](#footnote-ref-3)
4. See Pryor 1999. [↑](#footnote-ref-4)
5. See Morgan 2019 and Smith 2006. Smith imagines that the person with the pain, whenever they are in pain, presses a button that causes the other person to feel an itch. Obviously, there are lots of ways in which the causal connection could be implemented. [↑](#footnote-ref-5)
6. The *Simple Explanation* says that immunity to error through misidentification is explained by a judgment not being grounded in an identity, as opposed to by some claim about the nature of the self or the nature of mental contents. One challenge Coliva and Palmira identity (2024) for the *Simple Explanation* in their introduction is whether it can account for the distinction between logical and *de facto* IEM, which they treat as a desideratum on any explanation of IEM. To anticipate, my picture of memory involves denying that there is such a distinction. If this denial is correct, that desideratum disappears. Another challenge, which I don’t offer a response to but Coliva and Palmira do, is that the *Simple Explanation* struggles with wh-misidentification. For this criticism, see Merlo 2017. [↑](#footnote-ref-6)
7. See Morgan 2019, and McGlynn 2021. [↑](#footnote-ref-7)
8. ‘Shot in the dark’ is Evans’s phrase, used in his discussion of Shoemaker’s case, at (1982:246) [↑](#footnote-ref-8)