CHAPTER 2

Which 'key to all mythologies' about the self? A note on where the illusions of transcendence come from and how to resist them

Annalisa Coliva

It is a striking feature of philosophical reflection on the self that it often ends up being revisionary of our commonsensical intuition that it is identical to a living human being with, intrinsically, physical and psychological properties. As is well known, Descartes identified the self with a mental entity, Hume denied the existence of such an entity and Kant reduced it to a transcendental ego – a mere condition of possibility for experience and thought. In the *Tractatus*, Wittgenstein followed Kant – or, at any rate, the Kant made available to him through reading Schopenhauer – then, later, denied the existence of such an entity and proposed the no-reference view about at least some uses of 'I'. Finally, Anscombe radicalised Wittgenstein's views and claimed that no use of 'I' is ever referential.

It must be acknowledged that, despite the oddity of these views, philosophers have always arrived at their respective positions on the nature of the self through rational reflection: being impressed with some allegedly special feature of the use of 'I' (either in speech or in thought), they have felt compelled to account for it by postulating a realm of super-entities (or non-entities) which could explain such seeming peculiarities. Confronted with this tradition of revisionary accounts of the self, at least some contemporary theorists are now approaching the issue with a diagnostic eye, trying to identify the features that have led philosophers to embrace such positions, with the aim of offering a better understanding of them that could 'give philosophy peace'. That is to say, that could make them compatible with the commonsensical view that selves are identical to living human beings and that 'I', either in speech or in thought, is a genuinely referential expression. So, for instance, Christopher Peacocke

'The Key to All Mythologies' is Mr Casaubon's monumental, never-ending, basically pointless work aimed at uncovering the alleged kernel of truth common to all mythologies in George Eliot's *Middlemarch*, I, vii; III, xxix; IV, xlvii. I would like to thank Akeel Bilgrami, Pierre Jacob, François Recanati, Elisabeth Pacherie, Carol Rovane, Crispin Wright and Walter Pedriali for comments on previous versions of this chapter.

opens his influential and thought-provoking 'Self-knowledge and illusions of transcendence'[1] with the following remarks:

> Philosophical problems about the self and the first person provide a salient illustration of the challenge of integrating the epistemology and the metaphysics of a domain. There has been a persistent impulse amongst thinkers about the self to postulate a transcendental subject of experience and thought. It is an impulse to which Kant, Schopenhauer, Husserl and the early Wittgenstein all yielded. The impulse results from a combination of genuine insight and genuine error. The insight consists in the appreciation that there is an Integration Challenge which calls for philosophical solution. The error consists in trying, in this domain, to achieve integration by postulating an exotic domain of the transcendent, rather than by revising and deepening one's epistemology. (p. 263)

Rightly, Peacocke points out that 'there is probably more agreement on the preceding description of the situation than there is on the correct positive solution to this instance of the Integration Challenge' (*ibid.*). Hence, he sets the agenda for further work in this domain:

> The first task in this area is to identify adequately and precisely the distinctive feature of some first person thoughts which has led to illusions of transcendence. Once the feature is properly identified, the next task is the explanation of the existence of this feature, and some elaboration of its significance. (*Ibid.*)

While I agree both with Peacocke's description of the situation and with the issues he puts on our agenda, I disagree with much he says in carrying out this project. In particular, Peacocke identifies what he calls 'representational independence' – viz. the fact that psychological self-ascriptions are not grounded in first-personal contents, yet they feature the first-person concept – as the 'key to all mythologies' about the self.[2] For example: 'I am thinking about Pythagoras's Theorem' is grounded in one's thought about Pythagoras's Theorem, not on 'a mental event or state whose representational content is that *one* is thinking about Pythagoras's Theorem'.[3] Peacocke then takes representational independence to

[1] Peacocke 1999: ch. 6. Among other theorists who pursue a diagnostic project, see Shoemaker 1996; Wright 1998. By contrast, Evans 1982 and McDowell 1998 use immunity to error through misidentification of non-psychological self-ascriptions to *argue* in favour of a non-reductionist, so-called 'animalist' conception of the self.

[2] So, for instance, Peacocke claims that representational independence is the key to the no-reference view about 'I' held by Wittgenstein (p. 290), as well as to Lichtenberg's view that Descartes' *cogito* should be replaced with 'There is thinking going on now' (p. 291), as well as to other reductionist views about the self such as Hume's and Locke's (pp. 291–2).

[3] Peacocke 1999: 267, emphasis mine. Similar considerations would hold for 'I see the phone is on the table', 'I remember that Russell was born in 1872', 'I fear that the motion will not be carried' (p. 266). But see fn. 17.

be a genuine phenomenon, which any account of first-person thoughts will have to deal with and, finally, he provides a possible explanation of how representational independence may be compatible with the rationality, from a subjective point of view, of the relevant *self*-ascriptions.

By contrast, it seems to me that there can be more than one key to the mythologies about the self[4] and here I will take time to expound on another one: namely, the distinction – originally introduced by Sydney Shoemaker – between *de facto* and *logical immunity to error through misidentification*. Such a discussion will allow us to cast doubt on the legitimacy of the notion of representationally independent mental states and, in turn, to propose a dilemma for Peacocke's account: either he holds on to representational independence but he can't explain the rationality, from the first-personal point of view, of self-ascriptions of occurrent mental states, nor why the reference rule for 'I', according to which 'I = the thinker of this thought', holds as a matter of conceptual necessity and is a priori known to us; or else, he explains the latter phenomena but forsakes the former.

Furthermore, I will claim that the distinction between de facto and logical immunity to error through misidentification will allow us to vindicate Wittgenstein's intuition – discarded by Peacocke and generally fallen into disrepute – that, notwithstanding some important qualifications which will be introduced in the following, at least some *psychological* self-ascriptions enjoy a special kind of immunity to error through misidentification, which sets them apart from all others. Such an *asymmetry* – I submit – can in its turn account for the illusions of transcendence about the self, at least in part. For, if it is conceivable that the body one is receiving information from be not one's own, and that the (quasi-) memories one is storing depended on someone else's past, while it is inconceivable that the mental states one is immediately aware of be someone else's, the temptation will arise of identifying the self with the owner of such mental properties only, as Descartes did, or else, in a Kantian fashion, with their condition of possibility. Reacting to these ideas and *misrepresenting* introspective awareness, one may then be

[4] In my 2003 I have exposed another such key, which I called 'the real guarantee'. That is, the idea that any competent use of the first-person pronoun (either in speech or in thought) is such that one can't fail to know that the person one is thinking about (or referring to) when one uses it is oneself. I have also argued at length that it played a prominent role in Anscombe's way of reaching the startling conclusion that 'I' is never a referring expression. Another key is, obviously, the fact that any use of 'I' is guaranteed of reference (although, notoriously, Evans 1982: 249–55, contested it). Such a feature of first-person thought seems paramount in explaining Descartes' position.

tempted to identify the self with a sequence of appropriately related mental states, as neo-Lockeans do nowadays;[5] or else, to embrace either the view that there is no bearer of such mental states, as Hume did; or to deny, with Wittgenstein and Anscombe, that 'I' is a genuinely referring expression. I will then show why, granted the asymmetry between de facto and logically immune self-ascriptions, no such revisionary conclusions should follow.

IMMUNITY TO ERROR THROUGH MISIDENTIFICATION

As is well known, Wittgenstein, in the *Blue Book*, identifies the source of the temptation to think that at least some uses of the first person don't refer in what has been subsequently called by Sydney Shoemaker (1968) 'immunity to error through misidentification'. Wittgenstein's idea was that *all and only* psychological self-ascriptions are such that, although one could be wrong about the property one is ascribing to oneself, one couldn't be wrong about the fact that it is indeed oneself who has (or seems to have) the property in question. So, for instance, one's self-ascription 'I am seeing a tree' could be wrong because, in fact, one is having a hallucination and, thus, isn't seeing a tree, but it can't be wrong because one has made a mistake about the fact that it is oneself who is (or seems to be) seeing a tree. Now, Wittgenstein thought that since misidentification is always possible when empirical objects are at stake, only a super entity, which could be transparently present to the mind, such as a Cartesian, transcendent ego, could have secured immunity to error through misidentification of psychological self-ascriptions. Recoiling from the idea of a Cartesian ego, he then claimed that psychological self-ascriptions are such that the occurrence of 'I' in them is not genuinely referential and, in fact, similar in its function to 'it' in 'It's raining'.[6]

Notoriously, Wittgenstein's claims have been subject to a number of criticisms. The key element behind them has been the shift of the attention from the *subject-matter* of the self-ascription – its being a psychological as opposed to a non-psychological one – to the *grounds* on which these self-ascriptions are made. So, for instance, Gareth Evans (1982, pp. 219–20) has claimed that also some present-tense psychological self-ascriptions can be liable to error through misidentification when they are based on inference. Hence, varying Evans's example to make the point

[5] See Parfit 1984.
[6] As already noted, Wittgenstein's position has been extended by Anscombe 1975 to all uses of 'I'.

clearer, it is perfectly conceivable that a subject's judgment 'I hate my sister' could be affected by error through misidentification if, for instance, it happened to be made in the following circumstances. A subject takes part in psychoanalytic group sessions; he hears the therapist say 'You hate your sister' while looking pretty much in his direction. He believes anything the therapist tells him and, therefore, forms the belief 'I hate my sister'. In this case, the grounds of the subject's judgment would be: 'I am the person referred to by the therapist', 'That person hates his sister', therefore 'I hate my sister'. Since his final judgment is grounded in a suitable identification component – viz. 'I am the person referred to by the therapist' – it can be affected by error through misidentification: after all, the person referred to by the therapist could have been the one sitting next to the subject who is making the judgment.

Furthermore, Shoemaker, Evans and, more recently, John McDowell and Crispin Wright have made a convincing case that also bodily self-ascriptions, based either on perception or on somatic proprioception, can be immune to error through misidentification, as well as past-tense self-ascriptions of physical and psychological properties based on occurrent first-personal memories. So, for instance, when a subject judges 'I am in front of a tree', or 'My hair is blowing in the wind', or 'I was in Scotland five years ago', or, finally, 'I was angry', either on the basis of his perceptions or of somatic proprioception or of his occurrent first-personal memories, he may be wrong about the properties he is attributing to himself, but he cannot be wrong about the fact that he himself is the person who has (or seems to have) them.

Finally, all these theorists have pointed out that also demonstrative and indexical judgments can be immune to error through misidentification and yet no temptation would arise of thinking that there is nothing, or only something extraordinary, to which they refer. Hence, 'That tree is in blossom', 'It's cold here', or 'It's raining now' are all such that, although there is no possibility of misidentifying the object, the place, or the time in question, no temptation arises of thinking that there is no object, place or time to which 'that tree', 'here' and 'now' respectively refer. Of course in the case of 'that tree' a subject could be wrong about the *kind* of object demonstratively referred to. Yet this would be a mistake of *classification*, not of identification. That is, the subject would be wrong about the *sortal identity* of the object but he would have not taken *another* object to be identical to the one he is currently perceiving. Moreover, in the case of demonstrative judgments, the object referred to could fail to exist, if, for instance, one were having a hallucination. Yet again, the subject wouldn't

be wrong because he has taken *another* object to be the one he is aiming to refer to. For, if there is no object in the first place, then, a fortiori, there is no object to be misidentified either.[7] Finally, as the case of here- and now-judgments makes clear, it is not even the combination of immunity to error through misidentification with the impossibility of reference failure that could lead to the view that 'I' does not refer. For here- and now-judgments would have both these features and yet there would be no temptation to think that 'here' doesn't refer to a physical place and that 'now' doesn't refer to an instant in objective time.

For all these reasons Peacocke (1999, pp. 269–70, 286–9) discards immunity to error through misidentification as the source, or, at least, as one of the possible sources of the illusions of transcendence. However, let me offer some considerations to resist this conclusion. First, historically, Wittgenstein himself identifies immunity to error through misidentification both as the source of the Cartesian view about the self and of his own position. Thus, it must be acknowledged that it was significantly operative at least in his way of reaching his own view.[8] Secondly, to point out that it isn't true that all and only psychological self-ascriptions are immune to error through misidentification (and come up with an explanation of why this is so) belongs more to the 'cure' of the 'disease' than to its diagnosis. For it is only by realising that immunity to error through misidentification is a feature also of bodily self-ascriptions, and by

[7] Notice that demonstrative judgments based on identification components – either as part of their grounds or of their background presuppositions (see the next section) – can be affected by error through misidentification. For a nice example involving two loudspeakers, where one judges 'That speaker has gone dead', pointing at the one that is actually working, see Campbell 1997: 69. By contrast, Campbell's subsequent example about a transparent chair, perceived against a yellow background, so that one forms the false judgment 'That chair is yellow', seems to me to be spurious and that it could be simply accounted for as a case of mispredication. That is to say, as a case where one misperceives the colour of the chair one is demonstratively thinking of because its transparency and its proximity to yellow objects make *it* look yellow. Be that as it may, the demonstrative judgments which are usually taken to be immune to error through misidentification are the ones that do not involve any identification component (either in their grounds or background presuppositions). Moreover, the following remarks about here- and now-judgments will suffice to show that immunity to error through misidentification is compatible with reference to places and instants in objective space and time, respectively.

[8] Wittgenstein 1958: 69: 'We feel then that in the cases in which "I" is used as subject, we don't use it because we recognise a particular person by his bodily characteristics; and this creates the illusion that we use this word to refer to something bodiless, which, however, has its seat in our body. In fact *this* seems to be the real ego, the one of which it was said, "Cogito, ergo sum".' And at p. 67: 'And now this way of stating our idea suggests itself: that it is impossible that in making the statement "I have toothache" I should have mistaken another person for myself, as it is to moan with pain by mistake, having mistaken someone else for me. To say, "I have pain" is no more statement *about* a particular person than moaning is.'

providing an explanation of why this is so, that we can see that such an immunity is compatible with the fact that the relevant self-ascriptions can genuinely be about a living human being.[9] Finally, I think there is a further, deeper reason why immunity to error through misidentification cannot easily be discarded while trying to account for the illusions of transcendence. Namely, although it has been widely remarked that self-ascriptions based on somatic proprioception, perception and first-personal memories are also immune to error through misidentification, it has somehow been overlooked that, as Shoemaker (1986) once noticed, they are only *de facto* so. By contrast, *psychological* self-ascriptions based on occurrent sensations, perceptions, emotions, rememberings[10] and thoughts are *logically* so. Now, by saying that this *asymmetry* has somehow been overlooked, I don't wish to suggest that philosophers have failed to remark or acknowledge it.[11] Rather, what I mean is that they have systematically failed to provide an *adequate* explanation of it and to appreciate its *philosophical significance*. Hence, for instance, Peacocke himself (1999, pp. 286–9) recognises that asymmetry but then concludes that since logical immunity to error through misidentification is compatible with a genuinely referential use of 'I' – as the comparison with the case of demonstrative judgments shows – it can't be a genuine source of the illusions of transcendence and, correctively, of the no-reference view. As we will see, however, this judgment is based on failing to appreciate the deep philosophical significance of this asymmetry. Moreover, the reason offered to discard logical immunity to error through misidentification as one of the keys to the mythologies about the self seems to depend, once more, on the conflation between the diagnosis of the problem and its cure. For it is only when we see that demonstrative and indexical judgments can be so immune and yet make genuine reference that we will be able to reconcile the existence of this phenomenon with a non-revisionist conception of the self and of the role of 'I' in the relevant judgments.

[9] Notoriously, the cure is to argue that there is immunity because there is *no identification* going on in the first place, where 'identification' is understood in the thick sense of involving an identity judgment. Roughly: while there is identification in the sense of an individuation of a human being – oneself – which one refers to by means of 'I', there is no identification in the sense of an identification of a person with oneself, viz. 'I = that person/the such-and-such'. For a fuller treatment of the distinction between individuation and identification, see Evans 1982: 179–91 and ch. 7 (although he doesn't make use of the label 'individuation' and simply talks of the various immediate – that is, unmediated by any identity judgment – ways in which we can gain and store information about ourselves which allow us to have the first-person concept). See also my 2001, ch. 4.

[10] Hence, self-ascriptions such as 'I *remember* that *p*' based on one's occurrent memories and not self-ascriptions of past properties based on those memories' possibly first-personal contents.

[11] See Evans 1982: ch. 7; McDowell 1998: 372, and Peacocke 1999: 269–70.

DE FACTO IMMUNITY TO ERROR THROUGH MISIDENTIFICATION

The notion of de facto immunity to error through misidentification has been introduced by Sydney Shoemaker (1986) in the following way:

> [In the case of] 'I was angry' [said on the basis of memory in the ordinary way], a mistake of identification is impossible. It goes with this that [. . .] past-tense, first-person judgment does not rest on observationally based re-identification of the person referred to with 'I' [. . .] [But] one might 'quasi-remember' past experiences or actions that are not one's own [. . .] To allow that this is possible is to allow that *in a certain sense first-person memory judgments are subject to error through misidentification.* (p. 15, emphasis mine)

Evans (1982) then defined quasi-memories as follows:

> A subject quasi-remembers an event *e* iff (i) he has an apparent memory of such an event and (ii) that apparent memory in fact embodies information deriving from the perception of that event by a person who is not necessarily himself. (pp. 247–8)

So, we are usually asked to imagine cases of (partial) brain transplant which may not be physically possible, yet – all parties agree – are metaphysically so. Since quasi-memories are subjectively indistinguishable from ordinary memories, it follows that self-ascriptions of past properties based on the latter are in principle open to error through misidentification, although *contingently* not actually affected by it.[12]

The problem is to explain why, despite the fact that the relevant judgments aren't arrived at by holding in place any identification component of the form 'I am identical to the person whose past is responsible for the memory impressions I am now having', they are only contingently immune to error through misidentification. To put it otherwise: what is it about these judgments that makes it the case that, although they don't depend on an identification of the subject, they remain open to the *possibility* – in deviant scenarios – of error through misidentification?

Elsewhere,[13] I have proposed to account for this problem by deepening our understanding of the epistemological structure underlying these judgments. Accordingly, I have proposed to distinguish between an occurrence of an identification component as part of *a subject's own grounds* for a

[12] Similar considerations could be made in the case of bodily self-ascriptions based either on perception or on somatic proprioception.

[13] See my 2006.

given judgment, or as part of the *background presuppositions* of it. The idea, in very general terms, is this. In the former case, a body of information, while not necessarily part of a conscious inference, may be what a (rational and appropriately conceptually equipped) subject would appeal to if he were requested to justify his final judgment. By contrast, in the latter case, that very information, while neither entertained in a process of conscious inference, nor part of what the subject would appeal to in order to justify his judgment, may be such that, were it to be called into question or doubted, a (rational and conceptually equipped) subject ought to be prepared to withdraw from his final judgment.

This distinction has applications in various domains. Think, for instance, of visual perception and of judgments about material objects made on that basis. Consider, for example, the judgment 'There is a pole in front of me', made on the basis of my current visual experience. My ground for that judgment would be my visual experience as of a pole in front of me. Yet, this provides a warrant for 'There is a pole in front of me' just in case it is not called into question that there is a material world or that my sense organs are working reliably. 'There is a material world' and 'My sense organs are working reliably', however, certainly don't, nor need to feature as part of my own grounds – as my own avowable reasons – for my final judgment 'There is a pole in front of me'. Yet they are the background presuppositions on which my original judgment rests. For, if they were somehow to fall into question, I ought (rationally) to suspend from 'There is a pole in front of me'.[14]

Similarly, think of the recognition-based judgment 'John is wearing a white shirt'. If I were asked to justify my judgment I would offer the following grounds: 'That person is wearing a white shirt' and 'That person is John'. Still, it is obvious that that very judgment also rests on the background presupposition that John is immediately visually recognisable. Surely, however, I am not required to appeal to that, in giving my grounds for my judgment. Yet, were it to turn out that John

[14] It is then a further issue whether in order for my final judgment to be warranted those presuppositions should in turn be warranted – as a sceptic would maintain – or not. For opposite views on this issue, see Wright 2002, Pryor 2000. For a 'third way' which is alternative to both Wright's and Pryor's see my 2011. But notice that a sceptic could agree that *a particular subject* need not be able to appeal to those presuppositions, while offering his grounds for his judgment, and still insist that for the judgment to be warranted those presuppositions ought, in general, to be warranted. The disagreement between Wright, Pryor and myself hinges on the latter point: while Wright sides with the sceptic, Pryor thinks that it is merely sufficient that they aren't called into question or doubted, while I hold that they must be positively assumed.

has an identical twin brother who is in town and wears pretty much the same clothes, then I ought (rationally) to withdraw from my initial judgment.

With this distinction in hand, we can now account for the fact that memory-based self-ascriptions of past properties, while not grounded on any identification component, are nevertheless only *contingently* immune to error through misidentification. For, although an identification component such as 'I am identical to the person whose past is responsible for the memory impressions I am now having' is not and need not be part of the subject's own grounds for his memory-based I-judgments, it is nevertheless part of their background presuppositions. So, we may now say either that the judgment is immune to error through misidentification *relative to the subject's own grounds*, yet open to misidentification *relative to its background presuppositions*, thereby distinguishing between *two* different *kinds* of error and immunity to error through misidentification. Or else, we may say that there is just one kind of error through misidentification – viz. relative to one's own grounds – and hold that in abnormal metaphysical *and epistemic* conditions the relevant identification component could be moved from the background to the subject's own grounds for his judgment, and thus make the latter liable to error through misidentification after all. For instance, if we gave the subject the information that he might be storing memory impressions deriving from someone else's past *prior* to making his past-tense self-ascriptions, he could then offer the following grounds for his judgment: 'I seem to remember being in Scotland five years ago', 'I am the person whose past is responsible for the memory impressions I am now having', hence 'I was in Scotland five years ago', where the identification component could be wrong and the final judgment affected by error through misidentification (relative to the subject's own grounds).[15] Either way, the crucial point is that an identification component is being allowed to be part of the background presuppositions of one's judgment. For, no matter whether we take a liberal or a more conservative view on error through misidentification – allowing for two kinds of it, or just for one – it is the presence of such an identification component in the background presuppositions of one's judgment that allows us to account for

[15] Notice, however, that the role of the identification component, contrary to usual cases of error through misidentification, is not to ground an otherwise ungrounded *self*-ascription. Rather, its role is to ground the passage from the self-ascription of the logically weaker property of having *memory impressions* of being *F*, to the self-ascription of the logically stronger property of *having been F*.

the widely shared intuition that the relevant past-tense self-ascriptions are only *contingently* immune to error through misidentification.[16]

LOGICAL IMMUNITY TO ERROR THROUGH MISIDENTIFICATION

Shoemaker (1968) introduced the notion of logical immunity to error through misidentification as follows:

> 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*. (p. 89)

Shoemaker's point seems to be this. Contrary to non-psychological self-ascriptions and to past-tense self-ascriptions that are only contingently immune to error through misidentification, present-tense psychological ones, based on occurrent mental states, are so in all possible situations. This claim, in turn, appears to be based on two different considerations. First, on the *phenomenological* observation that in being *aware* that the attribute *feels pain* is instantiated, one is *ipso facto aware* of *oneself* instantiating it. Secondly, on the *conceptual* (or logical) point that being aware of feeling pain suffices for making that pain *one's own*. So, for instance, if, *per impossibile*, one's feeling pain depended on someone else's brain activity (let alone bodily affections), that awareness would suffice to make that pain one's own. Hence, on Shoemaker's view of the matter, there seems to be neither a phenomenological nor a logical gap between being aware that the attribute feels pain is instantiated, and being aware that *one oneself* is instantiating it.

Let us focus on the phenomenological claim first. It seems to me that a bit of introspection would easily make plain that also when one is thinking 'This train is late' – viz. when one has on occurrence of that sentence in mind – one is aware not just of the content of that thought, or

[16] By contrast, Evans (1982: 186) explicitly denies that identification components of the relevant kind could be part of the background presuppositions of one's memory-based judgments. Yet he insists that these judgments are only de facto immune to error through misidentification. But it is unclear to me how he – or anyone following him on this – could account for this latter intuition once he has deprived himself of the means to address it. For, if there is no room for any identification component as part of the underlying epistemological structure of the relevant judgments, how could they be liable to error through *misidentification*? They could only be wrong because of a mistake of predication, but then there would be no genuine distinction between de facto and stronger forms of immunity to error through misidentification. Thus, if there is such a distinction – as Evans himself (1982: 220) seems to think there is, for he thinks that at least some psychological self-ascriptions are incorrigible – an identification component *must* be acknowledged somewhere in the epistemological structure underlying the judgment.

that there is an episode of thinking that content, but also of the fact that that episode of thinking is going on in oneself – that *one oneself* is thinking 'This train is late'. Similarly, when one is aware of, and attending to what one is perceiving – for instance, when one is aware of and attending to the fact that there are people sitting next to one on the train – one is not just aware of the layout of the world, as it were, nor just of the fact that that layout is somehow given, but rather of the fact that *one oneself* is presented with it; thus that one oneself is perceiving so-and-so. Finally, when one has an occurrent memory, e.g. that yesterday *a* said such-and-so, one is not simply aware of *a*'s words, or of the fact that such an event was witnessed. Rather, one is aware of *oneself* remembering it.

If all this is correct, then Shoemaker's observation, extended to the examples just considered, speaks directly against Peacocke's claim that in all these cases the subject wouldn't be manifest to himself. To clarify: Peacocke is surely right in thinking that the subject isn't – or, at any rate, need not be – *re*presented in the *content* of these mental states. Yet, it doesn't follow that the subject isn't manifest to himself either in the act of performing a certain mental action, such as thinking that *p*; or else, as the recipient of a certain content given either through perception or memory. Hence, in effect, Peacocke's notion of representational independence seems to originate in a *non sequitur*: from the fact that in all these cases the self isn't represented as part of the content of certain mental episodes – i.e. from the fact that it isn't represented *as an object* – it doesn't follow that it isn't presented to itself *as a subject* – viz. as either the *agent* of certain mental activities like thinking a certain content, or else, as the *subject* of certain mental episodes, like perceiving or remembering that P, with their own distinctive contents.

It is important to notice that even if sometimes the phenomenology is passive, because, for instance, a subject is aware of himself as having a perception or an emotion that occurs to him and isn't subject to his own will, it remains that he is thereby aware of *himself* enjoying that specific mental state – even though as its *recipient*, as it were, rather than as its agent. Further to illustrate: just as it doesn't follow that one isn't (and isn't presented to oneself as) the subject who is either sending or receiving an envelope containing pictures, simply because one isn't (or need not be) the represented object of those pictures, so it doesn't follow that one isn't (and isn't presented to oneself as) the subject who thinks, feels, perceives and remembers, because one isn't (or may not be) represented as an object as part of the content of those mental episodes. Moreover, it is true that when one is presented to oneself as a subject one is not presented to

oneself through one's physical properties, still that doesn't stand in the way of being presented to oneself through one's own psychological attributes: as an agent or as a recipient of mental episodes with their own distinctive contents.[17]

Let me further stress that being presented to oneself through one's psychological affections does not consist in having a second-order thought with content 'I am thinking/perceiving/remembering that P'. Nor does it amount to having a perception of oneself as a mental substance (or of one's bodily self if the mental state in question had representational properties, like pain on some understanding of it, such as Evans's or indeed Shoemaker's). Rather, it must be taken as a further form of awareness which manifests a subject to himself through his current psychological affections. I don't think this is particularly mysterious. For, by being introspectively aware of the content of one's thoughts, one is also and *ipso facto* manifested to oneself as a subject capable of having introspective access to them. Obviously, being manifest to oneself this way – viz. as a subject – need not involve the possession of concepts such as 'I', 'introspective access', 'thought' and so on. Hence, probably the best way of thinking of the notion of being *presented* to oneself *as a subject* is to invoke a nonconceptual representation of oneself as enjoying conscious thoughts. But we can further enrich the picture. For, if I think 'This train is late', that thought doesn't contain any representation of myself as part of its content. Yet, it can originate in my previous deliberation to check whether we are travelling on time. Hence, it can be experienced as the immediate result of my own doings as a mental agent. Furthermore, the fact that *I* am introspectively aware of it can cause *me* to feel irritated – with the characteristic bodily affections that will entail – or can prompt *me* actively to engage in further thinking, which may be about *myself*, such as 'I'll be late for the meeting. In order to avoid it, I shall take a taxi. Let's hope it's already outside the railway station'. Hence, by enjoying these mental states one will also be manifest to oneself as a mental agent and as the subject of sensations and emotions, as well as a subject of conscious thoughts. Again, none of this seems to involve the exercise of concepts such as 'I', 'mental agent', etc., in the awareness one has of oneself as a subject. All that is required is to form a nonconceptual representation of oneself, which will then ground, in creatures endowed

[17] Peacocke 2008: 279–82 now recognises that at least when one is *judging* a content one is immediately aware of *oneself* as a mental agent, but he still thinks that other self-ascriptions of psychological properties depend on representationally independent mental states (p. 214).

with the concept 'I' and with all the relevant psychological concepts, the corresponding self-ascriptions.

Notice, moreover, that such a sequence of mental states at no stage involves the mere awareness of the content of the thought, the sensible wondering whether it is one's own, its further self-ascription and the subsequent arising of a sense of irritation and of further thoughts about oneself. Rather, that thought is immediately experienced as one's own – at least in virtue of its being introspectively available to one, if not also because of its being experienced as a result of one's own mental agency – and this reflects in the fact that it immediately gives rise to first-personal feelings and further thoughts that may be about oneself.

Similarly, when I look around and I notice that the other people sitting on the train are all smiling and relaxed, that perception, which doesn't have me as part of its content (bar perhaps the tip of my nose, or the frame of my specs), is immediately experienced as *mine* – at least in virtue of its content being introspectively available to me – and can then result in my sudden calming down and in my thinking 'Never mind, I'll be just a little late, it's all right', without going through any second-order thought in which one identifies it correctly as one's own perception. Finally, the memory of what *a* said yesterday, i.e. 'I'll be in my office all day', does not contain a representation of myself as an object. Yet, its being immediately given to *me* – because I am introspectively aware of its content – may further result in my suddenly feeling totally reassured and in thinking 'I can walk to the appointment after all, for *a* will be there all day long'. Again, none of this depends on having a certain memory content in view, on sensibly wondering whether it is my memory, on correctly self-ascribing it and on subsequently acting on it.

Hence, the fact that we have thoughts, perceptions and memories, whose contents we are conscious of, and which, moreover, can originate immediately in our mental activities, or passively affect us, yet cause us immediately to have other passive mental states, with distinctive first-personal contents, or to engage in further thoughts and deliberations that may be about ourselves captures – at least to some extent – the idea of being presented to ourselves through our psychological properties – or as subjects – that I've been trying to vindicate so far. Where, to repeat, none of this depends on having second-order thoughts about ourselves thinking/perceiving/remembering that P, possibly mediated by identification components of the form 'I = the thinker (or the subject) of this (introspectively available) thought (or further mental episode)'; or else, on having a perception of ourselves as mental entities, or (just) through our bodily affections.

Hence, this way, the Kantian 'I think' which must accompany all my psychological states neither is reified in a mental or in a physical substance, nor does it remain a mere transcendental entity, i.e. a mere condition of possibility for experience and thought, nor is it a second-order judgment accompanying otherwise representationally independent mental states. Rather, it gets grounded in some distinctive aspects of our phenomenology when we enjoy various kinds of psychological properties.

Now, as we have seen, self-ascriptions based on the presentation of the self *as an object* may either be liable to error through misidentification – when an identification component figures as part of the epistemic grounds on which they are based – or *de facto* immune to it, when that component is merely part of their background presuppositions. Moreover, we have seen that in both cases they may involve the ascription of either physical or psychological properties. In contrast, self-ascriptions based on a presentation of the self *as subject* are logically immune to error through misidentification and involve the attribution of *psychological* properties *only*.

As we saw before, from a conceptual point of view, their logical immunity would be due to the fact that, even if, *per impossibile*, those thoughts, sensations, perceptions and rememberings depended on someone else's brain activity, the mere fact that one is aware of them would be enough to make them one's own. So, to exemplify with a couple of cases: even if the pain one were feeling originated in someone else's brain state, the mere fact of feeling it (whether or not located in a specific part of one's body) would make it one's own. Similarly, even if the perception of people sitting next to one, or the thought 'This train is late', depended on someone else's brain activity, the simple fact of being aware of that perception or of that thought respectively would suffice to make them one's own. It merits note, moreover, that it is the phenomenology of introspective awareness that grounds this conceptual claim. That is to say, since in being aware of any given mental episode one is aware of *oneself* enjoying it, that suffices to make it one's own.

So much seems intuitively plausible. A problem arises, however, when we want to understand logical immunity to error though misidentification on the basis of the template proposed in the previous section. For we seem to have two options: either to say that psychological self-ascriptions based on the presentation of the self as subject are logically immune to error though misidentification because they don't contain any identification component as part of their background presuppositions; or else, because they contain such an identification component, which, however,

is a priori true and can't thus give rise to error through misidentification, not even in abnormal physical conditions.[18] As we have already pointed out, the phenomenology of mental experience does not speak in favour of identification components present in one's available grounds for a given self-ascription. Still, one might argue that, though this is a feature of our ordinary or normal experience, logical immunity to error through misidentification of psychological self-ascriptions should be explained by mobilising an a priori true identification component, operative in the epistemic background of our psychological self-ascriptions. Let us therefore look at this option in more detail.

The obvious candidate for the role of a priori true identification component would be 'I = the thinker (or, more generally, the subject) of this (presently given) thought (or sensation/perception/emotion/memory)'.[19] However, let me table some considerations to be suspicious of this strategy. Consider the following genuine identification components, which, *ex hypothesi*, we will take as true:

a) Hesperus = Phosphorus
b) I = the person whose reflection I am seeing in that window
c) I = the person whose past is responsible for the memory impressions I am now having.

In all these cases, the second term of the identification component presents the same object as the first, yet in a different way. Hence, if someone doubted whether (a)–(c) are true, they wouldn't be irrational (though, *ex hypothesi*, they would be wrong). In contrast, if someone doubted:

a) I = the thinker (or the subject) of this (presently given) thought (or sensation/perception/emotion/memory)

they would be irrational.[20] Now, such a verdict can't be explained by saying that (d) is *metaphysically necessary*, yet known a posteriori. For also

[18] Notice, however, that a revision of the definition of immunity to error through misidentification will then be needed. For, usually, it is defined in terms of the *absence* of any identification component (either in the background presuppositions, or in the subject's own grounds), whereas here it will depend on the presence in the background presuppositions of an identification component that is a priori true (cf. also fn. 9).

[19] Indeed, Peacocke seems to think that logical immunity to error through misidentification of psychological self-ascriptions of, in his view, representation-independent mental states, should be explained by reference to such an a priori true identification component. We will come back to Peacocke's position in the next section.

[20] Campbell 1999b maintains the opposite. For a reply, see my 2002a, 2002b. I personally think that the phenomenon of thought insertion on which Campbell draws in order to make his point is

(a)–(c), if true, would be metaphysically necessary truths, albeit known a posteriori. Rather, it could only be explained by saying that (d) is *analytic* – or *conceptually* necessary – and known a priori. But why would it be analytic and known a priori? The most likely explanation seems to me to be this: because there is no way of being presented with a thought without, *ipso facto*, being presented with *the subject* who is thinking it – with 'the thinker' mentioned on the right hand-side of (d) – and, in turn, without being *eo ipso* presented with *oneself*. Or, to put it differently, since in being aware of a given thought one is in fact aware of oneself thinking it, the identity between oneself and the thinker of that presently given thought is both analytic and known a priori.[21] In particular, what may be relevant and worth-pointing out is that the a priori knowledge of this identity isn't due merely to a stipulation, or to some kind of linguistic usage. Rather, it is grounded in the phenomenology of conscious thought I have tried to characterise in this section. Hence, here we are witnessing, I think, the possibility of conceptual and a priori truths, that are in fact grounded in some basic phenomenological features of our experience.[22]

Be that as it may, if in being aware of a given thought, one is *ipso facto* aware of oneself thinking it, the alleged identification component in (d) is not really such. That is to say, (d) doesn't encode the identity of a differently presented object with itself – it doesn't encode the identity of a still unidentified thinker of a presently given thought with oneself. Rather, it simply makes explicit the sense of the term on the left hand-side of the identity, viz. 'I' (equivalently, it merely individuates the first person concept). Hence, if (d) isn't a genuine identification component, it can't figure *as such* as part of the background presuppositions of self-ascriptions of psychological properties based on the presentation of the self as a subject. But if no genuine identification component does or can

merely the symptom of the fact that the subject enjoys a *passive* phenomenology not just when he is having sensations, perceptions, emotions and rememberings but also thoughts. Yet, as we saw before in the text, one can still be presented to oneself as a subject even if one is *passively* enjoying a given mental episode with its own characteristic content. So if, as I think, the distinction between being presented with oneself as a subject or as an object doesn't coincide with the distinction between being presented with oneself either as an active or as a passive subject, then the phenomenon of thought insertion does nothing to show that in those cases a subject wouldn't be presented to oneself as subject and could thus rationally disavow one's own introspectively available mental states.

[21] Similar considerations would be apposite for all other mental states mentioned in (d). For ease of exposition I have omitted further qualifications in the main text.

[22] Another candidate, as is well known, is the conceptual and a priori truth that an object can't be of two different colours all over its surface, at the same time, which would be grounded in the basic way in which we do experience colours. We need not discuss this case here.

indeed figure as part of those background presuppositions, then of course the resulting judgment will be *logically* immune to error through misidentification and not merely *de facto* so.

Let me stress, however, that, in my view, (d) is fine as a characterisation of the *sense* of 'I'. Yet, from the fact that (d) correctly individuates the first-person concept, it doesn't follow that it is a genuine identification component, which can figure as part of the background presuppositions of psychological self-ascriptions based on an awareness of oneself as subject. No more than saying that (e) – 'here = the place where I am' – states the identity of the same portion of space, yet given in two different ways, would follow from the fact that (e) correctly individuates the sense of 'here'.[23] Indeed, the fact that (d) and (e) specify the relevant *concepts excludes* that they may play the role of genuine identification components. For, if they just make explicit the concept they aim to characterize, it follows that one cannot rationally apply the concepts 'I' or 'here' (in 'I/here am/it is *F*') while sensibly wondering whether it is the thinker of a presently given thought that is *F*, or whether it is where one is that it is *F* (or vice versa). If one did, that would then automatically show either that one is being irrational[24] or else that one didn't really have those concepts in the first place. However, characteristically, the identification components which figure as part of the epistemic structure of judgments liable to some form of error through misidentification are such that their being incorrectly held true could neither automatically be seen as a case of irrationality nor, failing that, as a case of lack of the concepts involved.

So the point I have been making is really about the correct order of explanation and its consequences. On the view under consideration logical immunity to error through misidentification of some psychological self-ascriptions – based on representationally independent mental states – should be explained by appealing to the a priori truth of an alleged identification component. In contrast, on the view I am recommending, it is logical immunity to error through misidentification – due to the fact that in being aware of a thought one is presented with the subject of that mental episode and is *ipso facto* presented with *oneself* – that explains the a priori truth of (d) and allows it to play its proper role, viz. that of merely

[23] Similarly for the case of the concept 'now' which can be individuated, roughly, by 'now = the time at which I am thinking this thought'. If all this is correct, we can then explain why also genuine here- and now-judgments would be logically immune to error through misidentification. For they aren't based on any identification component, as the only possible candidate is in fact a concept-individuating identity. I can't, however, pursue this point here.

[24] That is precisely what the phenomenon of thought insertion would show, in my view. Cf. fn. 20.

individuating the first-person concept, rather than of being a genuine identification component somewhat operative in the epistemic structure underlying the relevant psychological self-ascriptions. Hence, I have claimed that proper consideration of the phenomenology of mental experiences speaks against Peacocke's notion of representational-independence and can be appealed to in order properly to account for logical immunity to error through misidentification and, connectedly, to individuate the first-person concept.

A DILEMMA FOR PEACOCKE'S ACCOUNT OF THE RATIONALITY OF PSYCHOLOGICAL SELF-ASCRIPTIONS

Before turning to a discussion of the philosophical implications of logical immunity to error through misidentification, let me briefly dwell on Peacocke's account to signal a problem with it, which the preceding considerations, if correct, will allow us to see more clearly. Recall that Peacocke claims that psychological self-ascriptions are grounded in the mere awareness of a mental state, which, as such, does not manifest its owner. So, for instance, my judgment 'I am thinking about Pythagoras' Theorem' is based on being aware of a thought about that theorem where such an awareness in no way manifests a subject of that mental episode.[25] On this view, even supposing for the sake of argument that it were supported by phenomenological considerations, the real difficulty would be to explain the *rationality* of the self-ascription: if no subject is manifested, how can the *self*-ascription of the mental state be rational? Peacocke's answer is that the self-ascription is rational because the transition from the representational-independent mental state to its self-ascription is made in accordance with the possession-conditions of the first-person concept, which, in turn, require one to be prepared to employ it when one is introspectively aware of a given mental state. As is well-known, according to Peacocke, concepts are individuated by their possession conditions. In particular, on his view, the first-person concept is individuated by the identity 'I = the thinker of this (introspectively available) thought'.[26] Peacocke's idea is that one's possession of the first-person concept is grounded in the disposition to use it when one is aware of a given thought – an awareness which, as such, doesn't manifest its owner – for the conceptual role of 'I' is precisely that of identifying the thinker of that introspectively

[25] Cf. fn. 3, 17.
[26] What Peacocke (2008: ch. 3) now calls 'the fundamental rule of reference' for 'I'.

available thought with oneself.[27] So, if one is aware of a given mental state, one is rationally entitled to self-ascribe it, because, even if such an awareness does not manifest oneself as the owner of the mental state, nor, in effect, does it manifest any subject, or 'thinker', of that mental state,[28] it is constitutive of having the first-person concept that one be prepared to *self*-ascribe that mental state on the basis of being introspectively aware of it.[29]

However, the problem for a proponent of such a strategy is to motivate the *rationality from a first-personal point of view* – so long as one shares Peacocke's internalist proclivities in epistemology – of that self-ascription.[30] A first, general worry is that, to such an end, it doesn't seem enough simply to say that a subject's transition from having a representationally independent mental state to its self-ascription conforms to an *externally* given rule, no matter how a priori and analytic we observers take it to be. Nor would it be wise to suggest that a subject's transition would be rational from the first-person point of view if explicitly conducted on the basis of his internalisation of that rule. For, clearly, not many would be able to make psychological self-ascriptions if they had to infer them by explicitly entertaining 'I = the thinker of this introspectively available thought'. Hence, the most promising explanation of the rationality, from a first-personal point of view, of present-tense psychological self-ascriptions based on occurrent thoughts would be to say that they are based on a subject's immediate and non-conceptual awareness of *himself* as thinking the thoughts he is aware of. As a consequence, the reference rule for 'I' isn't something a subject need consciously be able to entertain, in order to conform to it. Rather, it is an abstract specification of the first-person concept theorists may offer *ex post*. Yet, we could see subjects as conforming to it not merely because that allows us to make sense of their performance from a third-personal point of view, but because their use of 'I' would be grounded in their non-conceptual awareness of themselves as subjects, afforded by their

[27] Contrary to Peacocke (2008: 89–92), I think the conceptual role of 'I' should be liberalised a bit to include also introspectively available sensations, perceptions, rememberings and emotions. Be that as it may, I will omit this qualification in the following.

[28] This feature of one's introspective experience, on Peacocke's understanding of it, may in fact represent a further problem for his view. For if no subject of thought is presented to one, not even in a third-personal way, what would possibly ground the reference rule for 'I'? I will not pursue the matter here.

[29] See fn. 17.

[30] Such an explanation is needed, I think, even within the conceptual framework recently developed by Peacocke (2008) because the fundamental rule of reference for 'I' should also explain the norms that govern that concept and should provide a subject with internal reasons – that is, with subjectively available reasons – to move from mental states with certain contents to their self-ascription.

having conscious thoughts and further mental states. But now, clearly, such an account would be precluded if one followed Peacocke in holding that our inner phenomenology of thoughts, perceptions and memories consisted only in having representationally *independent* mental states.

A second, perhaps more serious worry is that if one endorsed a more externalistically oriented line and considered it enough that the transition rule be analytic and a priori known to *us*, yet applied on the basis of mental states which are representationally independent, one should still explain why that rule is analytic and a priori known to *us*. As we saw in the previous section, an overwhelmingly plausible account of those features would depend on the fact that in being aware of the content of an occurrent thought one would *eo ipso* be aware of *one oneself* thinking it.[31] Hence, even a more externalistically oriented strategy would face a similar problem. For to account for why the reference rule for 'I' is analytic and a priori known to us, we should appeal to the phenomenological datum that in being presented with a thought (or any other mental episode) one isn't merely presented with a given content, but, in fact, with oneself thinking it (or with oneself as being the recipient of a given perception, memory, etc.).

So, here is a dilemma for Peacocke. Either he insists on representational-independence as a genuine phenomenon, but he could neither motivate the rationality, from a first-personal point of view, of introspection-based psychological self-ascriptions, nor explain why 'I = the thinker of this thought' holds as a matter of conceptual necessity and is a priori known to us, or else he should maintain that representational independence is a pseudo-phenomenon, which a mature account of first-person thought shouldn't try, constructively, to account for and should rather denounce as brought about by a misleading conception of what introspective awareness amounts to. If the former, he should then join Lichtenberg in claiming that the only rationally sound judgment would be 'There is thought going on now'. If the latter, in contrast, he should provide an alternative explanation of the rationality of first-personal psychological ascriptions, which capitalizes on this phenomenological datum.[32]

[31] Notice that the appeal to the mere metaphysical necessity of the identity between the self and the thinker of a presently given thought, which seems to be Peacocke's new line (2008), could not provide such an explanation, as we saw in the previous section.

[32] I have put forward, though by no means developed, the suggestion that the self-ascription would be rational because it be based on one's being nonconceptually presented to oneself as a subject and would manifest the possession of the first-person concept, which would in its turn consist in being disposed to apply it on the basis of such a nonconceptual representation of oneself.

I, obviously, recommend doing the latter, for it seems to me quite plausible that phenomenological considerations in effect go against Peacocke's idea of representationally independent mental states and that arguments can show why 'I = the thinker of this thought' cannot be a genuine identification component. Still, it is important to clarify what kind of conception of introspective awareness may lead one to be genuinely impressed with the appearance of representational-independence. It seems to me that it is only when one thinks of mental states as *objects* presented to the subject in his mental arena that one can be seduced into thinking that representational-independence is a genuine phenomenon that stands in need of explanation. For, if in being introspectively aware of a mental state, what one is aware of is just a thought, a perception, or a memory that P, it may well seem mysterious and worthy of philosophical explanation how one could ever be rationally entitled to their *self*-ascription. Preoccupations with representational-independence, which, according to Peacocke, have led astray Hume, Wittgenstein, Kant, Lichtenberg and others, if and when they have been operative, are, therefore, the vestige of a way of thinking about introspection that is deeply Cartesian. If I am right, however, Peacocke himself is not immune to the seduction of that model and that is why he is profoundly impressed with representational-independence and tries to reconstruct the rational path which should lead us from there to the corresponding self-ascriptions. The problem is that, if the previous considerations are correct, the reconstruction can't work if one accepts its starting point – viz. representational-independence. The corrective will be to think of introspective awareness differently and thus to place crucial emphasis on the fact that introspective awareness is awareness of events which originate or happen to a subject who has an awareness of himself as a mental *agent* (or *recipient*). That is to say, who is aware of himself *as and through his enjoying* – actively or passively – various mental states. Hence, explaining the rationality, from a first-personal point of view, of introspection-based psychological self-ascriptions in effect requires giving up the idea of representational-independence.

A LITTLE QUIETIST CODA

If the above is correct we can finally see why introspection-based, psychological self-ascriptions are *logically* immune to error through misidentification. They are so because they do not contain as part of their background presuppositions any identification component, being, rather,

unmediated expressions of the awareness subjects have of *themselves* as enjoying mental states with a given content, where such an awareness suffices to make those mental states their own.

But now, if the distinction between de facto and logical immunity to error through misidentification is in good standing, we should ponder a bit on its philosophical significance. Essentially, what it amounts to is this: while it is readily conceivable that the body one is receiving perceptual and proprioceptive information from be not one's own, and that the (quasi-) memories one is storing depended on someone else's past, it is not conceivable that the sensations, perceptions, emotions, rememberings and thoughts one is immediately aware of be not one's own. That is to say, even if they could (metaphysically) originate in someone else's brain activity, their being objects of awareness of a given subject would make them his own. That – I submit – is why it remains a strong temptation to identify the self with the bearer of those psychological properties *only*. For while one's awareness of bodily, physical and past properties could, conceivably, depend on someone else's body being affected in various ways, or on someone else's past, one's occurrent awareness of psychological properties can't (logically) but depend on one's instantiating them. Hence, those latter properties might be thought to be the only essential properties we do have. This idea, in its turn, could (at least partially) motivate the notion that the self is just the bearer, or, at the minimum, the condition of possibility of psychological properties. Recoiling from the idea of such a disembodied yet self-standing self, and *misrepresenting* introspective awareness, one could then be tempted to embrace the view that the self is a mere construction out of representationally independent mental states, or else, that there is no such thing as the self and, connectedly, that 'I' isn't a referring expression at all.

However, if logical immunity to error through misidentification is one powerful source of the illusions of transcendence, we know, by now, what the cure is. The cure is to remind ourselves of the fact that immunity to error through misidentification of all stripes and sorts is perfectly compatible with genuine reference – as the case of demonstrative and of indexical judgments makes clear. Moreover, it is compatible with reference to an ordinary physical entity, which exists in space and time such as a human being. This point may escape notice because we have agreed that in being presented to oneself as a subject one is given to oneself only through one's psychological properties. That in turn might suggest the thought that logical immunity to error through misidentification can occur only when no physical entity is presented to one. That, however,

would be a mistake. For all logical immunity to error through misidentification entails is that there is no identification of a self operative in the epistemic background presuppositions of the relevant self-ascriptions. But of course this is entirely compatible with the fact that only a physically existing entity can enjoy mental states and be aware of itself through its own psychological affections. Hence, I have maintained that logical immunity to error through misidentification characterises only psychological self-ascriptions based on introspective awareness. Moreover, I have claimed that it ultimately depends on the fact that in introspection one is presented to oneself as the subject of psychological properties only. Yet neither of these claims entails that the self referred to in those self-ascriptions is or ought to be a mere mental entity, thus, acknowledging that the de facto/logical distinction is totally compatible with holding that in both cases the relevant judgments are about a living human being with physical and psychological properties, who can be presented to himself either through his physical attributes or through his psychological ones.

No doubt many will become impatient with this quietist conclusion, thinking that an 'animalist' conception of the self is a view that, if it can be maintained at all, can only be *earned* through philosophical spadework. It is the brief of this chapter that if one is such an 'animalist' then what one will be holding is, as a matter of fact, the *default* view. A view which, moreover, is entirely compatible with acknowledging the characteristic – and *prima facie* perplexing – features of self-reference and self-awareness, which may give rise to illusions of transcendence, such as logical immunity to error through misidentification of some psychological self-ascriptions, the impossibility of reference failure and what I have elsewhere dubbed 'the real guarantee'.[33] If so much is right, then, the burden of proof lies entirely on revisionists' shoulders.[34]

[33] See fn. 4.

[34] As remarked in fn. 1, McDowell, who is probably the most fervent quietist in philosophy nowadays, strikingly follows Evans in *arguing* from immunity of non-psychological self-ascriptions in favour of an animalist conception of the self and against Parfit's neo-Lockeanism. This attempt seems to me misguided on a number of fronts. Specifically because it isn't clear that Parfit's position can't accommodate this feature of memory-based self-ascriptions, as it has been convincingly argued for by Rovane 2006 and Buford 2009. A full diagnosis (let alone a 'cure') of what led Parfit to his revisionist position falls outside the scope of this chapter (but see Rovane 2006 for an analysis).