

Which Immunity to Error?*

A self-ascription is a thought or sentence in which a predicate is self-consciously ascribed to oneself. Self-ascriptions are best expressed using the first-person pronoun. Mental self-ascriptions are ascriptions to oneself of mental predicates (predicates that designate mental properties), non-mental self-ascriptions are ascriptions to oneself of non-mental predicates (predicates that designate non-mental properties). It is often claimed that there is a range of self-ascriptions that are immune to error through misidentification relative to the first-person pronoun (IEM for short). What this means, and exactly which self-ascriptions are properly classed as IEM, is a topic hotly disputed. Some claim that only mental self-ascriptions are IEM, others claim that some non-mental self-ascriptions are IEM. Before this question can be decided, it needs to be judged exactly what it means to say that a self-ascription is IEM. And here we stumble across the fact that there are, at least, two non-equivalent ways of defining the phenomenon¹. I will be claiming that one of these definitions should be rejected.

Before setting out my argument, it is worth considering what the point of the debate concerning IEM is. Standardly, the claim that a given class of self-ascriptions is IEM has been taken to suggest that those self-ascriptions are in some way particularly closely associated with self-consciousness. IEM is a feature of self-ascriptions that marks their centrality to our conception of ourselves as self-conscious subjects. Whence much of the interest in arguing for, or against, the claim that such-and-such a class of self-ascriptions (for instance, those based on memory, or bodily-awareness) are IEM. It follows from this that an account of IEM should, at the very

* Thanks to Lucy O'Brien, Mike Martin, Jim Pryor and Ann Whittle.

least, capture those self-ascriptions that are agreed by all to be central to our conception of ourselves as self-conscious subjects. Specifically, an account of IEM should imply that the self-ascription of occurrent mental episodes (e.g. ‘I have a headache’) are IEM. I argue that only one of the two definitions mentioned above can secure this claim.

1. The Standard Formulation of IEM

The phenomenon of IEM was first brought to our attention by Wittgenstein who, in *The Blue Book* claims that, “It is possible that, say in an accident, I should feel a pain in my arm, see a broken arm at my side, and think it is mine, when really it is my neighbours...On the other hand, there is no question of recognising a person when I say I have toothache. To ask ‘are you sure that it is *you* who have pains?’ would be nonsensical.” (Wittgenstein, 1958, pp.66-67). Wittgenstein is here pointing out that a certain class of self-ascriptions are immune to a particular kind of error².

As elucidated by Wittgenstein the notion of IEM is left rather vague. It is Shoemaker who first gives us a precise definition. He writes, “to say that a statement ‘*a* is ϕ ’ is subject to error through misidentification relative to the term *a* means that the following is possible: the speaker knows some particular thing to be ϕ , but makes the mistake of asserting ‘*a* is ϕ ’ because, and only because, he mistakenly thinks that the thing he knows to be ϕ is what ‘*a*’ refers to” (Shoemaker, 1968, pp.7-8).

One thing that clearly must be added to this definition is relativity to grounds. It is obvious that a judgement may be IEM when based on one kind of ground, but subject to error when based on another. To see this, compare introspection with

talking to one's analyst as ways of discovering one's mental states³. Through introspection it is impossible for me to discover that *someone* resents their mother, and mistakenly think that it is *I* who resent my mother. Through talking to one's analyst, on the other hand, it is quite possible for me to discover that *someone* resents their mother, and mistakenly think that it is me. So, judgements are IEM relative to the grounds upon which they are based.

According to Shoemaker's definition, the self-ascription 'I have a headache' is IEM relative to introspection, since it is impossible for me to know, via introspection, that some particular person has a headache but mistakenly think that it is I who have it. Coming to know, via introspection, that someone has a headache is nothing less than coming to know that I have a headache⁴. Shoemaker's explanation of how it is that such self-ascriptions are immune to error is that they are not based on an identification of the subject. It is impossible to misidentify oneself because there is no identification of oneself to go wrong.

This point is picked up by Evans in his treatment of IEM⁵. Immunity to error is a direct consequence of, what Evans calls, identification-freedom. According to Evans, to entertain a singular proposition about an object, the thinker must know which object the thought is about. One way this can be achieved is by means of an identification. For example, the thought '*a* is *F*' might be the result of the belief that '*b* is *F*' and the belief that '*a* = *b*'. Paraphrasing Evans, if knowledge of the truth of the singular proposition '*a* is *F*' can be seen as the result of knowledge of the truth of a pair of propositions '*b* is *F*' and '*a* = *b*' then we can call it 'identification-dependent'. The latter proposition '*a* = *b*' is its identification component. If a singular proposition is identification-dependent, the thinker knows which object it is about in

virtue of an identification. If, on the other hand, a thought does not rest on an identification component, then it is identification-free.

IEM, as defined by Shoemaker, follows from identification-freedom for the following reason: if a judgement is subject to error through misidentification, it must be possible for the thinker to make a mistake in the identification-component upon which it rests. But an identification-free judgement is not based on an identification-component, therefore it cannot be subject to error through misidentification, and so must be immune⁶. This close relationship between IEM and identification-freedom forms a common core running through the accounts of IEM given by Wittgenstein, Shoemaker and Evans. We can articulate this common core as follows⁷:

The proposition ‘ a is F ’ is immune to error relative to grounds G iff every possible judgement that a is F (based on grounds G) either:

(1) does not rest on two beliefs of the form ‘ b is F ’ and ‘ $a = b$ ’.

or

(2) rests on beliefs of the form ‘ b is F ’ and ‘ $a = b$ ’, but ‘ $a = b$ ’ is not rationally doubttable.

This definition captures the way in which IEM is a consequence of identification-freedom⁸. According to this definition, if a judgement is identification-free, when based on grounds G , it will be IEM. The necessity for condition (2) can be seen if we consider a judgement which is identification-dependent but which the identification upon which it is based is not rationally doubttable. It is plausible to think that this will

display immunity to error, since the indubitable character of the identification component will ensure that errors of misidentification will not occur. If this is right, identification-freedom and immunity to error could come apart.

Leaving such possible cases aside, we can think of IEM and identification-freedom as two sides of the same coin. We can also see how IEM is central to self-consciousness, in that it is intimately bound up with the functional role of self-ascriptions (what Evans calls 'I'-thoughts). It is a familiar point that 'I'-thoughts are liable to be affected by various special ways that we have of gaining knowledge about ourselves. These special ways clearly include introspection. Coming to believe, via introspection, that someone has a headache will dispose me to believe that I have a headache. It need not, however, dispose me to believe that Smith has a headache, even though I am Smith. The 'special ways' of gaining knowledge of oneself that play a part in the functional role of 'I'-thoughts just are those ways of gaining knowledge that lead to self-ascriptions that are IEM. So, arguing that self-ascriptions are IEM relative to a given faculty (say, bodily-awareness), is interesting for the reason that it helps us to give an account of the functional role of 'I'-thoughts.

2. Another Brand of Immunity to Error Through Misidentification?

Some philosophers have defined immunity to error through misidentification in a way significantly different from the account given above. For instance, Wright says that, "A claim made on a certain kind of ground involves immunity to error through misidentification just when its defeat is *not* consistent with retention of grounds for existential generalization" (Wright, 1998, p.19)⁹. This should be made clearer by

example. Suppose I judge ‘I was in Canterbury last week’, basing this upon my apparent memories of being in Canterbury. However, I am then informed that my apparent memory of being in Canterbury last week is, in fact, a quasi-memory. A quasi-memory being a state subjectively indistinguishable from a memory but which need not derive from the past history of the subject him or herself. If I have a quasi-memory of doing x , it may have been me that did x , or it may have been someone else. As such, memory is a sub-species of quasi-memory¹⁰. In this situation, I will no longer be justified in believing that *I* was in Canterbury last week, but I will be justified in making the general judgement ‘*Someone* was in Canterbury last week’. I can, as it were, retreat from the singular proposition to the general proposition, using the original grounds plus the piece of evidence that undercut my original judgement. The fact that I can do this, claims Wright, shows that my original judgement was not immune to error. On the other hand, consider the introspectively-based judgement ‘I have a headache’. If someone (incredibly!) undercuts my grounds for thinking that *I* have a headache, it is arguably not possible for me to retreat to the judgement ‘*someone* has a headache’. This shows that my judgement ‘I have a headache’ is immune to error.

Now all of this is compatible with the view that self-ascriptions based on memory are IEM as I defined it in the previous section. It may well be that memory-based self-ascriptions rest on no identity component, yet their grounds can be retained as grounds for existential generalization. This is enough to show that the kind of immunity to error which Wright et al have in mind really is something new. We might come up with another name for it then: for want of something more eloquent, I shall call it IEM*¹¹.

IEM* is a rare phenomenon. Much more so than its proponents have thought. Suppose that, on the normal perceptual grounds, I form the judgement 'I see a canary'. I am then informed that really, my visual impression as of a canary is hallucinatory, and is caused by someone else's seeing a canary. Whenever they see a canary they press a button which causes me to suffer a canary-illusion. In this case I will be justified in retreating to the existential proposition 'someone sees a canary'¹². Thus, there are at least some mental self-ascriptions that are not IEM*¹³.

What, then, is IEM* tracking if not identification-freedom? It is tempting to think that IEM* is the result of a judgement's being infallibly known or its being groundless. A judgement is IEM* when it is impossible for its grounds to be undercut in such a way as to leave open the possibility of the corresponding existential judgement. But it seems that for any judgement whatsoever, if it is possible to undercut its grounds, it is possible to undercut them in such a way as to leave open just such a possibility. We can formulate this as a general thesis: any self ascription 'I am *F*' which it is possible to undercut, can be undercut by the information that one suffers the illusion that one is *F* in just those cases when someone else is *F*. This allows one to retreat to the existential proposition '*someone* is *F*'. So IEM* arises in just those cases where a proposition's grounds cannot be undercut at all. This will turn out to be the case if the proposition is infallibly known, or if it is based on no grounds whatsoever.

But if IEM* just flags groundlessness or infallibility then it looks as if it defines too narrow a class of self-ascriptions to be of much independent interest. This is especially clear given that what many philosophers are looking to immunity to error to do is to capture that class of self-ascriptions seen as somehow central to self-

consciousness. The class of first-person thoughts which is central to self-consciousness has traditionally included self-ascriptions of experiences. But the proposition ‘I can see a canary’, as we have seen, is not IEM*. This might be conceded but it be pointed out that the judgement ‘I seem to see a canary’ is nevertheless IEM*. And it is judgements of *this* kind that are really expressive of self-consciousness. This seems fair enough. But as long as we agree that IEM* is tracking infallibility, there is scope for arguing that the introspectively-based self-ascriptions ‘I have a pain’ or ‘I am judging that P’ are not IEM*. This will be the case if, for example, it is possible to mistake an itch for a pain, or to mistake an imagining for a judging.

To spell this out in the case of ‘I am in pain’: suppose you experience what you take to be a pain. Now suppose someone reliably tells you that you are not really experiencing a pain but an itch, and that this has been caused by someone else’s suffering a real pain. Whenever they experience a pain, they press a button which causes you to feel a pain-like itch. In this case, you will be justified in retreating from the singular proposition ‘I have a pain’, to the general proposition ‘someone has a pain’ whilst retaining the same grounds. If this describes a possible situation, and it certainly seems to, then ‘I am in pain’ is not IEM*¹⁴. But, of course, there is little reason to think that the self-ascription of sensations is not somehow central to self-consciousness¹⁵.

My argument that IEM* is not a property of judgments such as ‘I am in pain’ may strike some as somewhat dubious. Of course, in these cases, the subject will be justified in retreating to the general proposition, but isn’t this justification all packed into the piece of defeating evidence? This feature might seem objectionable. Can’t we

simply reformulate our definition of IEM* so that the justification that the subject has for retreating to the general proposition must already be there in the justification for the original singular judgement? This would rule out my examples as counterexamples to the claim that fallible mental self-ascriptions can be IEM*. The problem with this suggestion is that it becomes impossible to distinguish IEM* from IEM. If, in the justification that I have for the singular judgement 'I am *F*' I am able to discern two elements, one of which can be discarded and the other retained, leaving me justified in thinking 'Someone is *F*', then it looks very much as though that original judgement was based on the information that *a* was *F* and that $I = a$, i.e. it was based on an identification¹⁶. And now the difference between the two forms of immunity to error through misidentification is entirely eroded. Thus, if the proponent of IEM* wishes that notion to be distinguished from the standard conception of IEM, he or she had better not make this change to the formulation.

I take it, then, that IEM* is not a particularly interesting phenomenon in its own right. Of course, the fact (if it is one) that some judgements are infallible or groundless is philosophically interesting, but we knew that already. It is misleading to treat IEM* as if it were a distinct phenomenon from infallibility¹⁷. I conclude, then, that if immunity to error through misidentification marks out those judgements that are closely bound up with self-consciousness, then it is IEM and not IEM* with which we should be concerned.

References:

- Campbell, J. 1999a. Immunity to Error through Misidentification and the Meaning of a Referring Term. *Philosophical Topics* 26:1&2 (1999), pp. 89-104.
Campbell, J. 1999b. Schizophrenia, the Space of Reasons, and Thinking as a Motor Process. *The Monist* 82:4 (1999), pp. 609-625.

- Campbell, J. 2002. *Reference and Consciousness*. Oxford: Clarendon Press.
- Evans, G. 1982. *The Varieties of Reference*, edited by J. McDowell. Oxford: Oxford University Press.
- Gallagher, S. 2000. Self-Reference and Schizophrenia. In D. Zahavi (ed.). *Exploring the Self: Philosophical and Psychopathological Perspectives on Self-Experience*. Amsterdam: John Benjamins Publishing Company, 2003.
- Hamilton, A. 1995. A New Look at Personal Identity. *The Philosophical Quarterly* 45:180, pp.332-349.
- Marcel, A. 2003. The Sense of Agency: Awareness and Ownership of Action. In J. Roessler and N. Eilan (eds.). *Agency and Self-Awareness: Issues in Philosophy and Psychology*. Oxford: Clarendon Press, 2003.
- Peacocke, C. 2003. Action: Awareness, Ownership and Knowledge. In J. Roessler and N. Eilan (eds.). *Agency and Self-Awareness: Issues in Philosophy and Psychology*. Oxford: Clarendon Press, 2003.
- Pryor, J. 1999. Immunity to Error through Misidentification. *Philosophical Topics* 26:1&2 (1999), pp. 271-304.
- Shoemaker, S. 1968. Self-Reference and Self-Awareness. *The Journal of Philosophy*, 65:19, pp. 555-67.
- Shoemaker, S. 1970. Persons and their Pasts. *American Philosophical Quarterly* 7:4, pp. 269-285.
- Strawson, P. 1966. *The Bounds of Sense: An Essay on Kant's Critique of Pure Reason* London: Methuen.
- Williamson, T. 2000. *Knowledge and its Limits*. Oxford: Oxford University Press.
- Wittgenstein, L. 1958. *The Blue and Brown Books*. Oxford: Blackwell.
- Wright, C. 1998. Self-Knowledge: The Wittgensteinian Legacy. In C. Wright, B. Smith & C. MacDonald (eds.), *Knowing Our Own Minds*. Oxford: Clarendon Press, 1998.

¹ As is made explicit in Pryor, 1999.

² Wittgenstein suggests that it is a misunderstanding of this *grammatical* feature of 'I' that tempts many towards the Cartesian view that "the real I lives in my body" (1958, p.66). A similar suggestion is made in Strawson, 1966, pp164-166.

³ By 'introspection' I do not intend a quasi-perceptual faculty. The term is simply a place-holder for that way, whatever it is, that we standardly come to know our own minds.

⁴ Campbell, 1999b, claims that the phenomenon of thought insertion shows that judgements based on introspection need not be IEM. This claim is sensibly disputed by Gallagher, 2000. A structurally similar debate concerning whether the phenomenon of anarchic hand syndrome shows that self-ascriptions of action based on proprioception are not IEM can be found in Marcel, 2003, and Peacocke, 2003.

⁵ Evans, 1982, Chs.6 & 7.

⁶ Pryor, 1999, argues that IEM does not follow from identification-freedom. This is due to his idiosyncratic definition of IEM.

⁷ My definition of IEM treats it as a property of propositions. However, throughout this section I have often spoken of individual judgements as IEM. Whilst this is a useful way of speaking, it may strike one as odd, since it is unclear in what sense individual judgements are *immune* to error – they are either errors or they are not. The claim that a judgement J is IEM should be understood as shorthand for the claim that (i) J expresses a proposition which is IEM relative to G, and (ii) J is based on G.

⁸ Evans' definition of IEM is actually slightly different. He defines it as follows, "a judgement is identification-free [and thus immune to error] if it is based upon a way of knowing about objects such that it does not make sense for the subject to utter 'Something is *F*, but is it *a* that is *F*?'", when the first component expresses knowledge which the subject does not think he has, or may have, gained in any other way." (1982, pp.189-190). This last constraint is not included in my definition of IEM. This should not worry us, since it is only necessary to introduce such a constraint if we rely on the idea of a judgement making sense for a subject. I avoid this locution, and intend Evans' and my definitions to be equivalent.

⁹ This different notion appears in discussions by Hamilton, 1995, Wright, 1998, Pryor, 1999, and Campbell, 1999a & 2002, Ch.5. In his, 1999b, Campbell appears to have something more like the standard definition of IEM in mind.

¹⁰ The notion of a quasi-memory was introduced by Shoemaker, 1970.

¹¹ Pryor, 1999, refers to this conception of immunity to error as 'immunity to *which*-misidentification'.

¹² This seems to me enough to cast very serious doubt on Pryor's suggestion that Shoemaker is primarily concerned with IEM*. See Pryor, 1999, pp.286-88; see Shoemaker, 1968, p.8, for his use of the canary example.

¹³ Wright, 1998, Campbell, 1999a, and Pryor, 1999, all assume that this sort of present-tense mental self-ascription is IEM*.

¹⁴ It may be suggested that this does not affect the definition of IEM* given by Pryor, 1999. He requires that the defeater be undercutting rather than additive (1999, p.284), and my defeater might be thought as a combination of undercutting and additive evidence. Although I can't argue it here, it is not clear to me that this distinction can bear the required weight.

¹⁵ If you are sceptical about the possibility of mistaking pains for itches etc. see the sorites-style argument in Williamson, 2000, Ch.4.

¹⁶ It has been suggested to me that the following is a counterexample to this claim: I think 'I have upset Joe', then due to defeating evidence retreat to 'Someone has upset Joe', but this does not show that the original judgement was based on an identity judgement. In order for this to be a counterexample it must not be a mixed case, that is the justification for the original judgement must derive entirely from a single faculty (say memory, or perception). Further it must be the case that although the original judgement was not based on an identity proposition, it is possible to defeat *part* of the original grounds, leaving the remainder to justify 'Someone has upset Joe'. What then does the original ground look like? It can't be the combination of 'Someone has upset Joe' and 'It was me' as that would be to return to the standard formulation of IEM, nor can it be simply 'I have upset Joe' for then it is completely unclear how a part of this can be defeated without the whole being defeated. In short, the purported counterexample relies on the thought that a *part* of my grounds can be defeated, but without explaining what the parts are. My contention is that one will have a hard job of showing that the parts do not contain an identity judgement.

¹⁷ This is somewhat unfair to Wright, 1998, who suggests that IEM* is not a particularly interesting phenomenon, since it is a trivial consequence of groundlessness. But whilst Wright correctly points out that groundlessness is sufficient for IEM*, the point of the present discussion is to argue that groundlessness (or infallibility) is both necessary and sufficient for IEM*.